

Fractional repetition codes with flexible repair from combinatorial designs

Oktay Olmez and Aditya Ramamoorthy

Abstract—Fractional repetition (FR) codes are a class of regenerating codes for distributed storage systems with an exact (table-based) repair process that is also uncoded, i.e., upon failure, a node is regenerated by simply downloading packets from the surviving nodes. In our work, we present constructions of FR codes based on Steiner systems and resolvable combinatorial designs such as affine geometries, Hadamard designs and mutually orthogonal Latin squares. The failure resilience of our codes can be varied in a simple manner. We construct codes with normalized repair bandwidth (β) strictly larger than one; these cannot be obtained trivially from codes with $\beta = 1$. Furthermore, we present the Kronecker product technique for generating new codes from existing ones and elaborate on their properties. FR codes with locality are those where the repair degree is smaller than the number of nodes contacted for reconstructing the stored file. For these codes we establish a tradeoff between the local repair property and failure resilience and construct codes that meet this tradeoff. Much of prior work only provided lower bounds on the FR code rate. In our work, for most of our constructions we determine the code rate for certain parameter ranges.

Index Terms—fractional repetition code, combinatorial design, Steiner systems, affine geometry, high girth, resolvable design, regenerating codes, local repair.

I. INTRODUCTION

Large scale data storage systems that are employed in social networks, video streaming websites and cloud storage are becoming increasingly popular. In these systems, the integrity of the stored data and the speed of the data access needs to be maintained even in the presence of unreliable storage nodes. This issue is typically handled by introducing redundancy in the storage system, through the usage of replication and/or erasure coding. However, the large scale, distributed nature of the systems under consideration introduces another issue. Namely, if a given storage node fails, it need to be regenerated so that the new system continues to have the properties of the original system. It is of course desirable to perform this regeneration in a distributed manner and optimize performance metrics associated with the regeneration process. Firstly, one

would like to ensure that the regeneration process be fast. For this purpose we would like to minimize the data that needs to be downloaded from the surviving nodes. Moreover, we would like the surviving nodes and the new node to perform very little (ideally no) computation, as this also induces a substantial delay in the regeneration process that is comparable to the download time (since nowadays, memory access bandwidth is comparable to network bandwidth [1]). In addition, the regeneration induces a workload on the surviving storage nodes and it is desirable to perform the regeneration by connecting to a small number of nodes. Connecting to a small set of nodes also reduces the overall energy consumption of the system.

In recent years, codes which are designed to satisfy the needs of data storage systems have been the subject of much investigation and there is extensive literature on this topic. Depending upon the specific metrics that are optimized there are different requirements that the distributed storage system needs to satisfy. However, broadly speaking, all systems have the following general characteristics. A distributed storage system (henceforth abbreviated to DSS) consists of n storage nodes, each of which stores α packets (we use symbols and packets interchangeably). A given user, also referred to as the data collector needs to have the ability to reconstruct the stored file by contacting any k nodes; this is referred to as the maximum distance separability (MDS) property of the system. To ensure reliability in the system, the DSS also needs to repair a failed node. This is accomplished by contacting a set of d surviving nodes and downloading β packets from each of them for a total repair bandwidth of $\gamma = d\beta$ packets. Thus, the system has a repair degree of d , normalized repair bandwidth β and total repair bandwidth γ . The new DSS should continue to have the MDS property.

A simple technique for obtaining a DSS is to treat the file that needs to be stored as a set of symbols over a large enough finite field, generate encoded symbols by using an MDS code (such as a Reed-Solomon (RS) code) and then store each encoded symbol on a different storage node. It is well recognized that the drawback of this method is that upon failure of a given storage node, a large amount of data needs to be downloaded from the remaining storage nodes (equivalent to recreating the file). To address this issue, the technique of regenerating codes was developed in the work of Dimakis et al. [2]. In the framework of [2], the repair degree $d \geq k$ and the system needs to have the property that a failed node can be repaired from *any* set of d surviving nodes. The principal idea of regenerating codes is to use subpacketization. In particular, one treats a given physical block as consisting of multiple

This work was supported in part by the NSF under grants CCF-1320416, CCF-1149860, CCF-1116322 and DMS-1120597 and by TUBITAK project numbers 114F246 and 115F064. The material in this work has appeared in part at the 50th Annual Allerton Conference on Communication, Control and Computing, 2012, the 2013 International Symposium on Network Coding and the 2013 Asilomar Conference on Signals, Systems and Computers.

Oktay Olmez (oolmez@ankara.edu.tr) is with the Department of Mathematics at Ankara University, Tandogan, Ankara, Turkey. Aditya Ramamoorthy (adityar@iastate.edu) is with the Department of Electrical and Computer Engineering at Iowa State University, Ames, IA 50011. Copyright (c) 2014 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

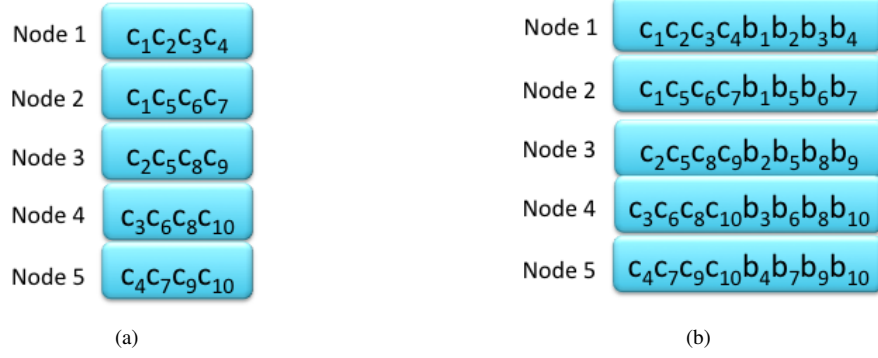


Fig. 1: (a) A DSS with $(n, k, d, \alpha) = (5, 3, 4, 4)$. Each node contains a subset of size 4 of the packets from $\{c_1, \dots, c_{10}\}$. First node for instance contains symbols $c_i, i = 1, \dots, 4$ that is $\{c_1, c_2, c_3, c_4\}$. When there is no confusion we simply use the notation $c_1 c_2 c_3 c_4$ instead of the set notation $\{c_1, c_2, c_3, c_4\}$. (b) The DSS is constructed by applying a $(20, 18)$ -MDS code followed by the inner fractional repetition code shown in the figure. It is specified with parameters $(5, 3, 4, 8)$. When a node fails we contact the remaining four nodes and download two packets from each to repair the failed node. This DSS can be obtained from the $(5, 3, 4, 4)$ DSS on the left by trivial β -expansion, where $\beta = 2$.

symbols (unlike the MDS code that stores exactly one symbol in each node). Coding is now performed across the packets such that the file can be recovered by contacting a certain minimum number of nodes. In addition, one can regenerate a failed node by downloading appropriately coded data from the surviving nodes. The work of [2] identified a fundamental tradeoff between the amount of storage at each node and the amount of data downloaded for repairing a failed node under the mechanism of functional repair, where the new node is functionally equivalent to the failed node, though it may not be an exact copy of it. Two points on the curve deserve special mention and are arguably of the most interest from a practical perspective. The minimum bandwidth regenerating (MBR) point refers to the point where the repair bandwidth, γ is minimum. Likewise, the minimum storage regenerating (MSR) point refers to the point where the storage per node, α is minimum.

In a different line of work, it has been argued that repair bandwidth is not the only metric for evaluating the repair process. It has been observed that the number of nodes that are contacted for purposes of repair is also an important metric that needs to be considered. The model of [2], which enforces repair from any set of d surviving nodes requires d to be at least k . The notion of local repair was introduced in [3], [4], [5], and considers the design of DSS where $d < k$. However, one only requires that there is some set of d surviving nodes from which the repair can take place.

The majority of work in the design of codes for DSS considers *coded* repair where the surviving nodes and the new node need to compute linear combinations of the stored symbols for regeneration. It is well recognized that the read/write bandwidth of machines is comparable to the network bandwidth [1]. Thus, this process induces additional undesirable delays [6] in the repair process. The process can also be potentially memory intensive since the packets comprising the file are often very large (of the order of GB). Motivated by these issues, reference [7] considered the following variant of the DSS problem. The

DSS needs to satisfy the property of *exact* and *uncoded* repair, i.e., the regenerating node needs to produce an exact copy of the failed node by simply downloading packets from the surviving nodes. This allows the entire system to work without requiring any computation at the surviving nodes. In addition, they considered systems that are resilient to multiple (> 1) failures. However, the DSS only has the property that the repair can be conducted by contacting some set of d nodes, i.e., unlike the original setup, repair is not guaranteed by contacting any set of d nodes. This is reasonable as most practical systems operate via a table-based repair, where the new node is provided information on the set of surviving nodes that it needs to contact. The work of [7] proposed a construction whereby an outer MDS code is concatenated with an inner “fractional repetition” code that specifies the placement of the coded symbols on the storage nodes. The main challenge here is to design the inner fractional repetition (FR) code in a systematic manner.

In this work, we present several families of FR codes and analyze their properties. This paper is organized as follows. In Section II, we outline our precise problem formulation, elaborate on the related work in the literature and summarize the contributions of our work. We discuss our FR code constructions for the case when $d \geq k$ in Section III, and explain the Kronecker product technique in Section IV. The locally recoverable FR codes where $d < k$ are considered in Section V and Section VI outlines the conclusions and opportunities for future work.

II. BACKGROUND, RELATED WORK AND SUMMARY OF CONTRIBUTIONS

A DSS is specified by parameters (n, k, d, α) where n - number of storage nodes, k - the minimum number of nodes to be contacted for recovering the file, d - the number of nodes to be contacted in order to regenerate a failed node and α - the storage capacity. In case of repair, the new node downloads β packets from each surviving node, for a total of $\gamma = d\beta$

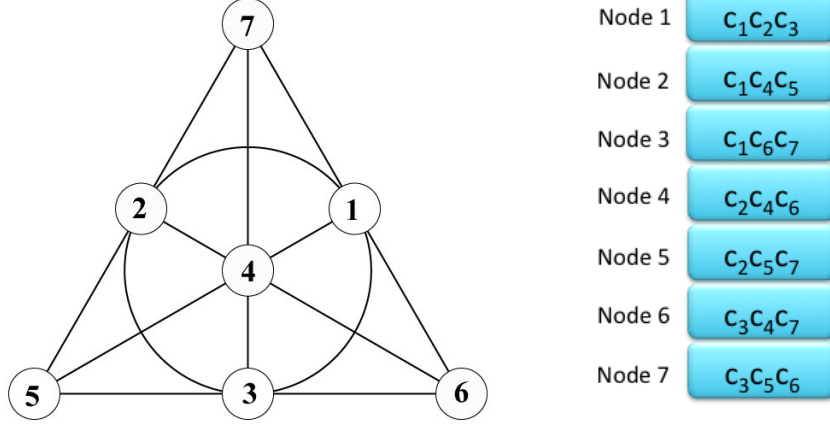


Fig. 2: $(7, 3, 3, 1)$ -BIBD also known as the Fano plane. Nodes of the DSS, which can be obtained from the Fano plane, are listed on the right.

packets. Let \mathcal{M} denote the size of file being stored on the DSS. We consider the design of fractional repetition codes that are best explained by means of the following example [8] with $(n, k, d, \alpha) = (5, 3, 4, 4)$.

Example 1: Consider a file of $\mathcal{M} = 9$ packets $(a_1, \dots, a_9) \in \mathbb{F}_q^9$ that needs to be stored on the DSS. We use a $(10, 9)$ MDS code that outputs 10 packets $c_i = a_i, i = 1, \dots, 9$ and $c_{10} = \sum_{i=1}^9 a_i$. The coded packets c_1, \dots, c_{10} are placed on $n = 5$ storage nodes as shown in Fig. 1a. This placement specifies the inner fractional repetition code. It can be observed that each c_i is repeated $\rho = 2$ times and the total number of symbols $\theta = 10$. Any user who contacts any $k = 3$ nodes can recover the file (using the MDS property). Moreover, a failed node can be regenerated by downloading one packet each from the four surviving nodes, i.e., $\beta = 1$ and $d = 4$, so that $\gamma = 4$.

Thus, the approach uses an MDS code to encode a file consisting of a certain number of symbols. Let θ denote the number of encoded symbols. Copies of these symbols are placed on the n nodes such that each symbol is repeated ρ times and each node contains α symbols. Moreover, if a given node fails, it can be exactly recovered by downloading β packets from some set of d surviving nodes, for a total repair bandwidth of $\gamma = d\beta$. It is to be noted that in this case $\alpha = \gamma$, i.e., these schemes operate at the MBR point. In the example above, $\beta = 1$, so that $\alpha = d$. One can also consider systems with $\beta > 1$ in general. A simple way to do this is replicating the symbols in the storage system. The resultant DSS has the parameters $(n, k, d, \beta\alpha)$ with $\beta > 1$. However, in this work we show that there are infinite families of FR codes with $\beta > 1$ which cannot be obtained this way. In Fig. 1b we illustrate the DSS obtained by replicating the $(5, 3, 4, 4)$ -DSS when $\beta = 2$.

Before introducing the formal definition of a fractional repetition (FR) code we need the notion of β -recoverability. Let $[n]$ denote the set $\{1, 2, \dots, n\}$.

Definition 1 (β -recoverability): Let $\Omega = [\theta]$ and $V_i, i = 1, \dots, d$ be subsets of Ω . Let $V = \{V_1, \dots, V_d\}$ and consider $A \subset \Omega$ with $|A| = d\beta$. We say that A is β -recoverable from V if there exist $B_i \subseteq V_i$ for each $i = 1, \dots, d$ such that

$$B_i \subset A, |B_i| = \beta \text{ and } \cup_{i=1}^d B_i = A.$$

Definition 2 (FR Codes): A fractional repetition (FR) code $\mathcal{C} = (\Omega, V)$ for a (n, k, d, α) -DSS with repetition degree ρ and normalized repair bandwidth $\beta = \alpha/d$ (α and β are positive integers) is a set of n subsets $V = \{V_1, \dots, V_n\}$ of a symbol set $\Omega = [\theta]$ with the following properties.

- (a) The cardinality of each V_i is α .
- (b) Each element of Ω is contained in exactly ρ sets in V .
- (c) Let V^{surv} denote any $(n - \tau)$ sized subset of V and $V^{fail} = V \setminus V^{surv}$. Each $V_j \in V^{fail}$ is β -recoverable from some d -sized subset of V^{surv} . Let ρ_{res} be the maximum value of τ such that this property holds.

We provide the following example to illustrate that requirement (c) of Definition 2 plays an important role in our study.

Example 2: Consider the sets $\Omega = \{1, 2, 3, 4, 5, 6\}$, and two different families of subsets of Ω as shown below.

$$V = \{\{1, 2, 3\}, \{2, 3, 4\}, \{4, 5, 6\}, \{1, 5, 6\}\}, \text{ and} \\ W = \{\{1, 2, 3\}, \{3, 4, 5\}, \{2, 5, 6\}, \{1, 4, 6\}\}.$$

Both V and W satisfy the requirements (a) and (b) of Definition 2. However, note that $\{1, 2, 3\} \cap \{4, 5, 6\} = \emptyset$. This implies that $\{1, 2, 3\}$ is not 1-recoverable from the set

$$\{\{2, 3, 4\}, \{4, 5, 6\}, \{1, 5, 6\}\}.$$

So $C = (\Omega, V)$ cannot be a fractional repetition code. In contrast, any failed set in W is 1-recoverable and thus $C = (\Omega, W)$ is a fractional repetition code with $\delta = 1$.

The value of ρ_{res} is a measure of the resilience of the system to node failures, under the constraint of exact and uncoded repair. The file size is given by

$$\mathcal{M} = \min_{I \subset [n], |I|=k} |\cup_{i \in I} V_i|$$

and the code rate is defined as $R_C = \frac{\mathcal{M}}{n\alpha}$. We emphasize that R_C depends on k .

Note that the parameters of a FR code are such that $\theta\rho = n\alpha$. Thus, the code rate $R_C = \frac{\mathcal{M}}{n\alpha} \leq \frac{\theta}{n\alpha} = \frac{1}{\rho}$. Moreover as

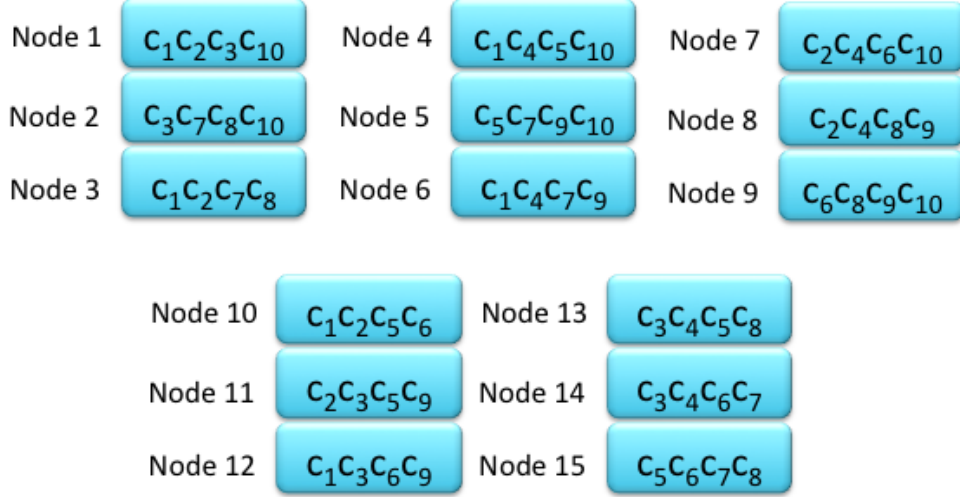


Fig. 3: The figure shows a DSS where $n = 15, k = 4, d = 2, \theta = 10, \alpha = 4, \rho = 6$. A node can be repaired by contacting the other two nodes in the same column. The system is resilient up to 5 node failures.

$\rho \geq 2$, the maximum rate of any FR code is at most $\frac{1}{2}$. It is to be noted that the parameter \mathcal{M} also sets the code rate of the outer MDS code; it is exactly \mathcal{M}/θ . For a FR code $\mathcal{C} = (\Omega, V)$ and an index set $\mathcal{I} \subseteq [n]$, we say that nodes $V_i \in V$ for $i \in \mathcal{I}$ cover ζ symbols if $\zeta = |\cup_{i \in \mathcal{I}} V_i|$.

The work of [7], only considered FR codes with $\beta = 1$ and $k \leq d$, i.e., for recovery the new node would contact d surviving nodes and download a single packet from each of them. For their codes, the requirement (c) in Definition 2 is satisfied and the system is resilient to $\rho - 1$ failures, i.e., $\rho_{res} = \rho - 1$. It is to be noted that the requirement of $d \geq k$ is essential in the problem formulation considered in [2] since the systems require node recovery from any set of d surviving nodes. In that setup if $d < k$, it is easy to see that one can always specify a failed node and a set of d nodes from which recovery is impossible. However, in the framework of [7], the recovery requirement is relaxed. Specifically, to recover from a failure, the new node contacts a specific set of nodes from which it regenerates the failed node. Thus, the recovery process is table-based and for each node we only need to guarantee the existence of one set of d nodes from which recovery is possible. Thus, it becomes possible to have systems with $d < k$. In fact, in Section V, of this paper, we present several constructions of FR codes where $d < k$. In the literature, these are referred to as codes that allow for local repair.

For FR codes, the failure resilience ρ_{res} and the code rate R_C are two evaluation metrics and it is evident that there is a tradeoff between them. Indeed, if the outer MDS code does not add any redundancy, i.e., $\mathcal{M} = \theta$ then k would need to be chosen such that any k nodes cover all the θ symbols and the code rate of the system would be exactly $\frac{\theta}{n\alpha}$. However, in this case the DSS will be resilient to at most $\rho - 1$ failures under any possible recovery procedure, i.e., even without any constraint

on the repair. In contrast, if the outer code introduces nontrivial redundancy, the file size \mathcal{M} would be lower but it may be possible to reconstruct the DSS in the presence of more than $\rho - 1$ failures. To see this, consider Example 1 where the outer MDS code has rate $9/10$. Note that under exact and uncoded repair, this DSS is resilient to only one failure. However, the DSS can be reconstructed even in the presence of the failure of any two nodes, since any three surviving nodes cover at least nine symbols. Our proposed codes will also be evaluated in terms of their *minimum distance* which quantifies this tradeoff.

Definition 3 (Minimum Distance of a DSS): The *minimum distance* of a DSS denoted d_{\min} is defined to be the size of the smallest subset of storage nodes whose failure guarantees that the file is not recoverable from the surviving nodes. The Singleton bound on the minimum distance in this context can be found, e.g., in eq. (15) in reference [9].

Lemma 1 (Singleton Bound): Consider a DSS with parameters (n, k, d, α) with file size \mathcal{M} and minimum distance d_{\min} . Then,

$$d_{\min} \leq n - \left\lceil \frac{\mathcal{M}}{\alpha} \right\rceil + 1.$$

It turns out that codes that have the local repair property, i.e., codes with $d < k$ suffer a penalty on the maximum possible minimum distance. This tradeoff was captured in the case of scalar (i.e., $\alpha = 1$) codes by [3] and by [4] in the case of vector (i.e., $\alpha > 1$) codes.

Lemma 2: Consider a DSS with parameters (n, k, d, α) with file size \mathcal{M} and minimum distance d_{\min} . Then,

$$d_{\min} \leq n - \left\lceil \frac{\mathcal{M}}{\alpha} \right\rceil - \left\lceil \frac{\mathcal{M}}{d\alpha} \right\rceil + 2.$$

We note that if $d \geq k$, we have $\lceil \frac{\mathcal{M}}{d\alpha} \rceil = 1$ so that the bound above reduces to the Singleton bound.

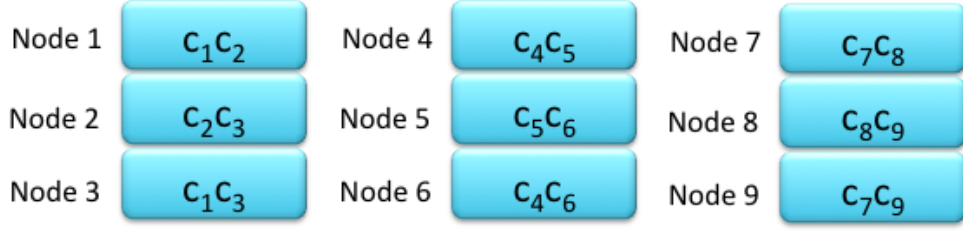


Fig. 4: A failed node can be recovered by contacting two nodes and downloading one packet from each. The code is resilient up to five failures and the file size is 5. The minimum distance is of the code is 6, since any four nodes can recover the file.

Observation 1: A given DSS meets the Singleton bound if $k = \lceil \frac{M}{\alpha} \rceil$. Similarly, a code meets the bound in Lemma 2 if $k = \lceil \frac{\rho}{\alpha} \rceil + \lceil \frac{M}{d\alpha} \rceil - 1$.

It is to be noted that the bound in Lemma 2 holds for all possible local repair codes. In this work, we consider the added constraint that the repair takes place purely by download. Thus, for our constructions, the bound in Lemma 2 is in general loose. In Section V we derive a tighter upper bound on the minimum distance of codes where the repair process is local and operates purely by download.

At various points we will need to use the well-known inclusion-exclusion principle for computing the maximum file sizes that can be supported by our DSS. For the sake of completeness, we state the result here.

Theorem 1: [Inclusion-Exclusion principle] Consider n sets A_1, A_2, \dots, A_n . If $\mathcal{I} \subseteq [n]$, let $A_{\mathcal{I}} = \cap_{j \in \mathcal{I}} A_j$. Then

$$|A_1 \cup A_2 \cup \dots \cup A_n| = \sum_{\emptyset \neq \mathcal{I} \subseteq [n]} (-1)^{|\mathcal{I}|+1} |A_{\mathcal{I}}|. \quad (1)$$

It can also be shown that

$$|A_1 \cup A_2 \cup \dots \cup A_n| \geq \sum_{i=1}^n |A_i| - \sum_{i < j} |A_i \cap A_j|. \quad (2)$$

Many of our constructions will result from combinatorial designs that we briefly introduce (a detailed description can be found in [12]).

Definition 4 (Combinatorial Design): A combinatorial design (or, simply a design) is a pair (Ω, V) where Ω is a finite set of elements called “points” and V is a collection of non-empty subsets of Ω called “blocks”.

A prototypical example with several applications is the balanced incomplete block design (BIBD).

Definition 5 (Balanced Incomplete Block Design): A $(\theta, \rho, \alpha, \lambda)$ balanced incomplete block design (BIBD) is a pair (Ω, V) that forms a combinatorial design such that $|\Omega| = \theta, |V| = n$; every element of Ω is contained in exactly ρ blocks and every 2-subset of Ω is contained in exactly λ blocks.

Let n denote the number of blocks. By using combinatorial double counting arguments it can be seen that for a BIBD, the following relations hold.

$$n\alpha = \theta\rho, \text{ and} \quad (3)$$

$$\rho(\alpha - 1) = \lambda(\theta - 1). \quad (4)$$

A $(\theta, \rho, \alpha, \lambda)$ -BIBD can be used as the FR code in a DSS as long as β -recoverability is guaranteed for an appropriate β (there are several instances when $\beta = 1$). These $(\theta, \rho, \alpha, \lambda)$ -BIBDs include finite projective planes and affine planes. Table II contains a list of well-known families of Steiner systems. A $(a^2 + a + 1, a + 1, a + 1, 1)$ -BIBD is equivalent to a *projective plane* of order a . Projective planes have interesting geometric properties that can be used in determining the corresponding file size. For instance, any two blocks of a $(a^2 + a + 1, a + 1, a + 1, 1)$ -BIBD share exactly one point and any two points are contained in exactly one block in a projective plane. The smallest example of a projective plane corresponding to $a = 2$ is known as the *Fano plane* and is depicted in Fig. 2. For more information on the projective planes and affine planes we refer the Chapter 2 of [12].

One can use the Fano plane to design the inner FR code, by interpreting the points as symbols and the blocks as storage nodes. Suppose we first apply a $(7, 6)$ -MDS to the file. Then we can place the coded symbols on the storage nodes as depicted in Fig. 2. Note that these storage nodes are obtained from the blocks. The obtained DSS has the property that any two nodes share exactly one symbol. Thus, using Theorem 1 contacting any three nodes recovers at least $9 - \binom{3}{2} = 6$ distinct symbols and hence the file. Furthermore, we can identify a set of three nodes whose intersection is empty, e.g., nodes 1, 2 and 4. Thus, the maximum file size this DSS can support is 6. An affine plane can be obtained by deleting one block and its all points from a projective plane. Hence, an *affine plane* of order a is equivalent to a $(a^2, a + 1, a, 1)$ -BIBD. Here any two points are contained in exactly one block. However, there are in general pairs of blocks that do not have any points in common. More generally, a FR code can be obtained from Steiner systems.

Definition 6 (Steiner Systems): A $S(t, \alpha, \theta)$ Steiner system is a set Ω of θ elements and a collection of subsets of Ω of size α called blocks such that any t -subset of the symbol set Ω appears exactly one of the blocks.

Steiner systems are examples of t -designs. A FR code is a t -design if every t -subset of symbols is contained in exactly λ nodes. The concept of t -designs can be viewed as a generalization of the concept of BIBDs. Naturally, a $S(2, \alpha, \theta)$ Steiner system is a $(\theta, \rho, \alpha, 1)$ -BIBD where

$$\rho = \frac{\theta - 1}{\alpha - 1}, \quad n = \frac{(\theta - 1)\theta}{(\alpha - 1)\alpha}, \quad d = \alpha \text{ and } \beta = 1.$$

Method	$(n, \theta, \alpha, \rho)$	\mathcal{M}	Range of k	Comments
Steiner Systems with $t = 2$	$(n, \theta, \alpha, \frac{\theta-1}{\alpha-1})$	$\geq k\alpha - \binom{k}{2}$	$1 \leq k \leq \alpha$	Steiner systems with $\alpha = 3, 4, 5$ are completely characterized and explicit constructions are known. Here we list the necessary and sufficient conditions for the cases $\alpha = 3, 4, 5$ <ul style="list-style-type: none"> Steiner systems with $\alpha = 3$ exists for any $\theta \equiv 1, 3 \pmod{6}$. Steiner systems with $\alpha = 4$ exists for any $\theta \equiv 1, 4 \pmod{12}$. Steiner systems with $\alpha = 5$ exists for any $\theta \equiv 1, 5 \pmod{20}$. For $\alpha \in \{6, 7, 8, 9\}$ there are only finitely many exceptions where the existence of Steiner systems is unknown. For this we refer the reader to the tables provided in section 3 of the book [10].
Transposed Steiner Systems	$(n, \theta, \frac{\theta-1}{\rho-1}, \rho)$	$\geq k\alpha - \binom{k}{2}$	$1 \leq k \leq \alpha$	File size equals $k\alpha - \binom{k}{2}$ if the original Steiner system has a maximal arc. <ul style="list-style-type: none"> There exist a Steiner system with $\alpha = 3$ and a maximal arc if $\theta \equiv 3, 7 \pmod{12}$. There exist a Steiner system with $\alpha = 4$ and a maximal arc if $\frac{\theta-1}{3}$ is a prime power. To our best knowledge results about the existence of maximal arcs in Steiner systems with higher values of α are not known.
Grids	$(2a, a^2, a, 2)$	$ka - k^2/4$ for even k , $ka - (k^2 - 1)/4$ for odd k	$1 \leq k \leq a$	The file size calculation can be done for any positive integer a .
MOLS (Remark 4)	$(4a, a^2, a, 4)$	$\geq k\alpha - \binom{k}{2}$	$1 \leq k \leq 4$	File size equals $k\alpha - \binom{k}{2}$ if $k = 4$. Use the construction of two MOLS for order greater than 6 [11].
MOLS (Lemma 9)	$(\rho p^m, p^{2m}, p^m, \rho \leq p^m - 1)$	$kp^m - \binom{k}{2}$	$1 \leq k \leq \rho$	p is a prime. Use the construction of MOLS where the order is a prime power.

TABLE I: Constructions where $d \geq k$ and $\beta = 1$. Note that we can perform trivial β -expansion to obtain higher β .

α	θ	Comments
q	q^2	q is a prime.
$q + 1$	$q^2 + q + 1$	q is a prime.
$q + 1$	$q^3 + 1$	q is a prime. These designs are known as Unitals.
2^r	$2^{r+s} + 2^r - 2^s$	$2 \leq r < s$. These designs are known as Denniston designs.

TABLE II: Well-known infinite families of Steiner systems when $t = 2$. These can be found in [10].

Thus, projective planes and affine planes are instances of Steiner systems. A given FR code can be put in one-to-one correspondence with an incidence matrix as explained below.

Definition 7 (Incidence Matrix of a FR Code): An incidence matrix of a FR code $\mathcal{C} = (\Omega, V)$ where $\Omega = [\theta]$ and $V = \{V_1, V_2, \dots, V_n\}$ is the $\theta \times n$ binary matrix N defined by

$$N_{i,j} = \begin{cases} 1, & \text{if } i \in V_j; \\ 0, & \text{otherwise.} \end{cases}$$

We shall sometimes refer to the FR code \mathcal{C} by simply referring to its incidence matrix N . We will occasionally refer to the bipartite graph corresponding to the FR code as well. This is defined next.

Definition 8 (Bipartite graph of a FR Code): For a FR code $\mathcal{C} = (\Omega, V)$ where $\Omega = [\theta]$ and $V = \{V_1, V_2, \dots, V_n\}$ with incidence matrix N , we define its bipartite graph $G_b =$

$(V_l \cup V_r, E)$ as follows. We associate the storage nodes in V with the vertices V_l and the points in Ω with the vertices V_r so that V_l and V_r are disjoint. There exists an edge between $v \in V_l$ and $u \in V_r$ if and only if $N(u, v) = 1$.

Definition 9 (Transposed FR Code): For a FR code \mathcal{C} with incidence matrix N , the code specified by N^T is called *transposed FR code* of \mathcal{C} and denoted by \mathcal{C}^T if the design obtained from N^T is β -recoverable for some β .

Note that, in the transposed code, the roles of the storage nodes and the symbols are reversed. An infinite family of transposed codes can be obtained from Steiner systems with $t = 2$. In such Steiner systems any pair of symbols is contained in exactly one node which implies that any pair of nodes in the transposed design share exactly one symbol. This in turn means that the transposed design is 1-recoverable.

Incidence matrices with appropriate parameters can be combined via operations such as the Kronecker product to

Method	$(n, \theta, \alpha, \rho)$	β	\mathcal{M}	Range of k	Comments
Affine Resolvable Designs	$(q\rho, q^m, q^{m-1}, 1 \leq \rho \leq \frac{q^m-1}{q-1})$	q^{m-2}	$q^m \left(1 - \left(1 - \frac{1}{q}\right)^k\right)$	$1 \leq k \leq m$	The file size exceeds the trivial lower bound $k\alpha - \beta \binom{k}{2}$. The parallel classes need to be chosen in a careful manner. If $q > m$, then we choose $\rho > m$ and if $q \leq m$, we choose $\rho \leq m$.
Hadamard Designs	$(8a-2, 4a, 2a, 4a-1)$	a	$3a$	$1 \leq k \leq 2$	β is not restricted to be a prime power.

TABLE III: Constructions where $d \geq k$ and $\beta > 1$. In many cases these construction parameters cannot be obtained by trivial β -expansion.

Base Code	Method	$(n, \theta, \alpha, \rho)$	β	\mathcal{M}	Range of k	Comments
\mathcal{C}^T obtained via the transpose of a Steiner system $S(2, \tilde{\alpha}, \tilde{\theta})$ with $\tilde{\rho} = \frac{\tilde{\theta}-1}{\tilde{\alpha}-1}$ with maximal arc of size $\tilde{\rho}+1$	Kronecker product of \mathcal{C}^T with itself	$(\tilde{n}^2, \tilde{\theta}^2, \tilde{\alpha}^2, \tilde{\rho}^2)$	$\tilde{\alpha}$	$k\tilde{\rho}^2 - \tilde{\rho} \binom{k}{2}$	$1 \leq k \leq \tilde{\rho}$	There exist a Steiner system with $\tilde{\alpha} = 3$ and a maximal arc if $\tilde{\theta} \equiv 3, 7 \pmod{12}$. In several cases, the codes obtained via Kronecker product cannot be obtained by trivial β -expansion.

TABLE IV: Constructions obtained via Kronecker product, where $d \geq k$.

obtain new matrices (equivalently FR codes) with a new set of parameters. We use this technique extensively in the sequel to generate families of FR codes.

Definition 10 (Kronecker Product): If A is an m -by- r matrix and B is a p -by- q matrix, then the *Kronecker product* $A \otimes B$ is the mp -by- $r q$ matrix

$$\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1r}B \\ a_{21}B & a_{22}B & \cdots & a_{2r}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mr}B \end{pmatrix}.$$

Let N_1 and N_2 be two incidence matrices of FR codes $\mathcal{C}_1 = (\Omega_1, V_1)$ and $\mathcal{C}_2 = (\Omega_2, V_2)$ with parameters $(n_1, \theta_1, \alpha_1, \rho_1)$ and $(n_2, \theta_2, \alpha_2, \rho_2)$ respectively. Let c_1, \dots, c_{n_1} be the n_1 columns of N_1 and d_1, \dots, d_{n_2} be the n_2 columns of N_2 . A new FR code can be obtained from the old one by the following incidence matrix

$$\bar{N} = [N_1 \otimes d_1 \quad N_1 \otimes d_2 \quad \cdots \quad N_1 \otimes d_{n_2}].$$

We can find an appropriate permutation matrix P such that the matrix $\bar{N}P$ is equal to the Kronecker product of N_1 and N_2 . Note that matrix P reorders the columns of \bar{N} .

We can obtain a DSS by replicating the symbols of another DSS via the Kronecker product. In the subsequent discussion we will refer to this technique for obtaining codes with $\beta > 1$ as *trivial β -expansion*.

Definition 11: [Trivial β -expansion] Let N be incidence matrix of a FR code \mathcal{C} with parameters $(n, \theta, \alpha, \rho)$ with $\beta = 1$. Let $\mathbf{1}$ be the $m \times 1$ all-ones column vector. The FR code $\tilde{\mathcal{C}}$ obtained from $\tilde{N} = N \otimes \mathbf{1}$ which has parameters $(n, \theta m, \alpha m, \rho)$ is called a trivial β -expansion of the code \mathcal{C} with $\beta = m$.

In the remainder of this section, we discuss some illustrative examples of FR codes. Our first example is a code with $\beta > 1$ that cannot be obtained by trivial β -expansion.

Example 3 (A Non-trivial Code with $\beta > 1$): Consider the DSS shown in Fig. 3. The ten symbols are obtained by using an outer $(10, 6)$ MDS code followed by the FR code illustrated in Fig. 3. Note that the DSS can recover from a single node failure by downloading two packets each from two nodes in the same column; hence $d = 2$. Moreover, any two nodes share 0, 1, or 2 symbols in common which implies that any two nodes recover at least 6 symbols, thus $k = 2$. According to the Singleton bound $d_{\min} \leq 15 - \lceil \frac{6}{4} \rceil + 1 = 14$. The system requires only two surviving nodes to recover the file thus the code is resilient up to 13 failures (since $k = 2$) and thus meets the Singleton bound. However, this code (with $\beta = 2$) cannot be arrived at simply by replication. To see this we note that if this were true, the original DSS with $\beta = 1$ must correspond to a storage capacity of 2 and have a number of symbols which is 5. However, this means that there can be at most $\binom{5}{2} = 10$ distinct storage nodes of capacity two. Thus our design with $n = 15$ cannot be obtained this way.

The idea underlying Example 3 can be formalized as follows.

Observation 2 (Non-trivial FR Codes with $\beta > 1$): A FR code with parameters $(n, \theta m, \alpha m, \rho)$, $\beta = m$ and distinct storage nodes cannot be obtained from a trivial β -expansion if $n > \binom{\theta}{\alpha}$.

Next, we demonstrate an example of a locally recoverable DSS, i.e., a system where $d < k$ that is constructed using the Kronecker product method.

Example 4 (Locally Recoverable Code Using Kronecker Product Technique): Let $\mathcal{C} = (\Omega, V)$ be a FR code with $\Omega = \{1, 2, 3\}$ and $V = \{V_1 = \{1, 2\}, V_2 = \{2, 3\}, V_3 = \{1, 3\}\}$ with incidence matrix N . The code obtained from $\tilde{N} = I \otimes N$

Method	$(n, \theta, \alpha, \rho)$	β	\mathcal{M}	Range of k	Comments
Use undirected graph $\Gamma = (V, E)$, $ V = n$, degree s and girth g	$(n, \frac{ns}{2}, s, 2)$	1	$k(s-1)$	$k = as + b$	Need $s > b \geq a + 1$ and $k \leq g$. Codes meet the minimum distance bound for locally recoverable codes. Since an (s, g) -cage minimizes the number nodes in the system, we will have highest possible code rate for this particular construction. Here we list some of the well-known infinite families of (s, g) -cage. <ul style="list-style-type: none"> • $(s, 3)$-cages are complete graphs on $s + 1$ vertices. • $(s, 4)$-cages are complete bipartite graphs on $2s$ vertices. • When $s - 1$ is a prime power $(s, 6)$-cages can be obtained from incidence graphs of projective planes. • When $s - 1$ is a prime power $(s, 8)$-cages and $(s, 12)$-cages can be obtained from incidence graphs of generalized polygons. For more see the survey [13].
Use l copies of FR code $\mathcal{C} = (\Omega, V)$ with parameters $(\tilde{n}, \tilde{\theta}, \tilde{\alpha}, \tilde{\rho})$ such that any $\tilde{\Delta} + 1$ nodes cover $\tilde{\theta}$ symbols. $ V_i \cap V_j \leq \tilde{\beta}$ for $i \neq j$. Parameters satisfy $(\tilde{\rho} - 1)\tilde{\alpha}\tilde{\theta} - (\tilde{\theta} + \tilde{\alpha})(\tilde{\Delta} - 1)\tilde{\beta} \geq 0$.	$(l\tilde{n}, l\tilde{\theta}, \tilde{\alpha}, \tilde{\rho})$	$\tilde{\beta}$	$t\tilde{\theta} + \tilde{\alpha}$	$t\tilde{n} + 1$	Codes meet the minimum distance bound for locally recoverable codes with exact and uncoded repair (cf. Section V).

TABLE V: Constructions of local FR codes where $d < k$.

is presented in Fig. 4 where I denotes the 3×3 identity matrix. Suppose that the outer MDS code has parameters $(9, 5)$, so that $\theta = 9, \mathcal{M} = 5$. Consider contacting any of the four nodes depicted in Fig. 4. These nodes will fall into one of the three columns in the figure. So, there are three cases we need to examine.

- **Case (a):** Two nodes can be chosen from one of the columns and one from each of the rest. The union of these nodes has a cardinality of 7.
- **Case (b):** We first select two columns and two nodes within each column. In this case the size of the union is 6.
- **Case (c):** Finally, we can select two columns and choose three nodes in one column and one node in the other column. In this case the cardinality of the union is 5.

Thus, it is evident that contacting any $k = 4$ nodes will recover at least 5 symbols. Note that a failed node can be recovered by contacting the remaining two nodes in its column by downloading one packet from each of them. Thus, $d = 2 < k$. This implies that the code is locally recoverable. By applying a similar case analysis for the failure patterns we can conclude that the code is resilient to 5 failures and it meets the minimum distance bound in Lemma 2.

A. Summary of Contributions

In this work we present several constructions of FR codes. The contributions of our work can be summarized as follows.

We construct a large class of FR codes for $d \geq k$ from combinatorial structures such as grids, mutually orthogonal Latin squares (MOLS), resolvable designs and Hadamard designs. These were first presented in the literature in the conference version of the current manuscript [14]. While [7] presented constructions based on Steiner systems, our work presents a rigorous analysis of the file size of the corresponding DSS. The Kronecker product technique for generating new DSS from existing ones is also new [15]. Furthermore, our conference paper [16] was the first to present locally recoverable FR codes where $d < k$.

Tables I – V contain a description of the various constructions and the corresponding DSS parameter values that can be achieved by these constructions. We defer an in-depth discussion of these parameters to the respective sections. However, we highlight the key contributions of our work by referring to appropriate rows of Tables I – V below. Specific details about the construction techniques can be found in the corresponding sections of the paper.

- We construct a large class of FR codes based on resolvable designs [12] where the repetition degree (ρ) of the symbols can be varied in an easy manner (see Table I (rows 3 – 5) and Table III). The constructions of [7] lack this flexibility as they are mostly based on Steiner systems where the repetition degree is usually fixed by the construction.
- We construct FR codes where $\beta > 1$, i.e., the new node

downloads more than one packet from the d surviving nodes. We emphasize that starting with a FR code with $\beta = 1$, it is trivially possible to arrive at a code with $\beta > 1$ by trivial β -expansion (*cf.* Definition 11). However, such a strategy only results in a limited range of system parameters that can be achieved. We present several codes (see Tables III and IV) that achieve certain parameter ranges that cannot be achieved in a trivial manner.

- Determining the file size that can be supported by a given FR code turns out to be challenging. Much of the literature in combinatorial designs only discusses the pairwise overlaps between the content of the different storage nodes. However, the file size depends on the union of all subsets of storage nodes of size k . In this work we determine the file sizes for most of our constructions. In particular, we demonstrate a family of FR codes whose file size is strictly larger than a simple lower bound that is obtained by applying the inclusion-exclusion principle (see row 1, Table III). We also determine the file size for a large class of codes obtained from Steiner systems that were originally considered in [7] (see row 2, Table I). Several of our constructions are shown to meet the Singleton bound for specific file sizes, which demonstrates their optimality.
- We present the Kronecker product as a technique for constructing new FR codes from existing ones (Table IV) and analyze the properties of codes thus obtained.
- In this work, we propose a large family of *locally recoverable* FR codes where $d < k$, i.e., the repair degree is strictly smaller than the number of nodes contacted for recovering the stored file. We derive an appropriate minimum distance bound for our class of codes that enjoy local, exact and uncoded repair, and demonstrate constructions that meet these bounds (Table V).

B. Discussion of related work

The work of Dimakis et al. [2] initiated the work on regenerating codes, by demonstrating the tradeoff between the storage capacity of nodes and the repair bandwidth. Their work considered functional repair, where the new node is functionally equivalent to the failed node and demonstrated that random network coding suffices for achieving this tradeoff. Following this, several papers [17], [8], [18], [19], [20], [21], [22], [7], [14] considered the construction of exact repair regenerating codes, where the new node is an exact copy of the failed node. In most cases, these constructions either operate at the minimum storage regenerating (MSR) point [17], [21], [23], [22], [18] or the minimum bandwidth regenerating (MBR) point [17], [8], [24], [7], [14]. More recently, codes with local repair have been investigated where the metric for repair is the number of surviving nodes that are contacted for repair [3], [5], [4], [9], [25], [16].

Constructions of repair-by-transfer codes, where node repair is performed simply by downloading symbols from surviving nodes was first presented in the work of [24] where they constructed a repair-by-transfer MBR code with $d = n - 1$. Repair by transfer codes have also appeared in [26], [27].

The work of [7] also considered such codes (termed “exact and uncoded repair”) but with a repair degree that can be strictly smaller than $n - 1$. The repair operates by contacting a specific set of d surviving nodes and is hence table based. Reference [7] introduced the system architecture whereby an MDS code is applied to a file consisting of \mathcal{M} symbols to obtain θ symbols. These symbols are then placed onto the storage nodes and this placement is referred to as the fractional repetition (FR) code. The codes in [7], were derived from Steiner systems. They provided lower and upper bounds on the corresponding file sizes. Following this, the work of [28] constructed FR codes from bipartite cages. These codes enjoy the property that the node storage capacity is much larger than the replication degree. For the given parameters they design codes with the smallest number of storage nodes. In [28], they used MOLS to construct bipartite cages and the codes thus obtained are different from ours. In our construction we obtain the storage nodes directly from the set of MOLS and also obtain net FR codes. Reference [29] presents necessary and sufficient conditions on the existence of a FR code with certain parameters; however, it does not consider the issue of determining the file size for a given k .

The work of [30] presents several FR code constructions based on combinatorial structures including regular and biregular graphs, graphs with a given girth, transversal designs, projective planes and generalized polygons. They consider codes where $\alpha = d \geq k$ and $\beta = 1$ and show that the file size of their constructions meets the upper bound presented in [7] for $k \leq d$. This work is closely related to the content of Section III of our work. Their construction of FR codes from transversal designs treats the blocks of the transversal design as symbols. Thus, it can be considered as working with the transpose of the incidence matrix corresponding to the original transversal design. Our FR codes in Section III are obtained from nets which can also be viewed as transposes of transversal designs. However, as discussed in Section III-C, the analysis of file size for our constructions cannot be obtained from the results in [30]. Our work differs in the sense that we present constructions with non-trivial β values, Kronecker product constructions and local FR codes.

The problem of local repair for scalar codes ($\alpha = 1$) was first considered in [3]. This was extended to vector codes ($\alpha > 1$) in [9], [4]. References [9], [4] study the tradeoff between locality and minimum distance and corresponding code constructions. In [9], the authors presented constructions that use the repair-by-transfer MBR codes of [24] as individual components. Local codes were also studied in [31] where the design consists of an outer Gabidulin encoder followed by inner local MBR encoders. This work (see Construction III.1 in [31]) also provides examples of local FR codes by using t -designs. However, the achievable parameters are limited as k needs to be chosen to be at most t and explicit constructions of t -designs for large t are largely unknown (when $t \geq 3$ there are only finitely many known explicit constructions [10]). In Section V we focus on regenerating codes that allow a repair process in a local manner by simply downloading packets from the surviving nodes. We provide an upper bound for the minimum distance and constructions of codes which meet

this bound. Our constructions use local FR codes instead of repair by transfer MBR codes. We also note that our codes are quite different from those that appear in [31], [9] and allow for a larger range of code parameters. Regenerating codes using t -designs were also presented in [19]. The architecture of the codes consists of a layered erasure correction structure that ensures a simple decoding process. These codes are showed to be achieve performance better than time-sharing between MBR and MSR points.

III. CONSTRUCTION OF FR CODES WHEN $k \leq d$

In this section we present the construction of FR codes where $d \geq k$. As discussed in Example 2 it is possible that certain set systems do not satisfy the property of β -recoverability and hence cannot be used to construct FR codes. However, there are a large class of combinatorial designs that can be used to construct FR codes. In particular, we present various constructions of FR codes that are derived from balanced incomplete block designs (BIBDs) and resolvable designs. Our constructions address several issues that exist with prior constructions in the literature. For instance, resolvable designs allow the repetition degree of the symbols in the FR code to be varied in a simple manner, a flexibility that prior constructions typically lack. We present a large class of codes that cannot be obtained via trivial β -expansion.

Our first set of constructions are FR codes based on Steiner systems with $t = 2$ (that are BIBDs) which have been previously considered in the literature [7]. However, to our best knowledge, prior work does not provide results on the file size of the constructions. In the discussion below, we present a certain class of Steiner systems for which we can determine the file size of the FR codes obtained from their transpose. To demonstrate the difficulty of determining the file size for a general Steiner system, we first discuss two non-isomorphic Steiner systems with the same parameter values that result in FR codes with different file sizes. This demonstrates that file size calculations for Steiner systems cannot be performed just based on the system parameters. Accordingly, we consider Steiner systems that have maximal arcs [32], [33]. It turns out that we can determine the file size of the corresponding transposed codes.

A. FR codes from Steiner systems

We consider Steiner systems $S(2, \alpha, \theta)$. Note that the repetition degree of any symbol is $\rho = \frac{\theta-1}{\alpha-1}$ and any two distinct symbols are contained in exactly one node. Consider the FR code $\mathcal{C} = (\Omega, V)$ obtained from it and its transpose.

In general, it is a challenging task to find the file size for a given FR code. For codes obtained from Steiner systems and their transposes, lower bounds based on the inclusion-exclusion principle were presented in [7]. However, it is important to note that the file size depends critically on the structure of the Steiner system, i.e., two Steiner systems with the same parameters can have different file sizes. To see this, consider two non-isomorphic Steiner systems $S(2, 3, 15)$ denoted \mathcal{D}_1 and \mathcal{D}_2 ; the nodes of these designs are provided in Tables VI and VII. These designs can also be found in [10].

{0, 1, 2}	{0, 3, 4}	{0, 5, 6}	{0, 8, 7}	{0, 9, 10}
{0, 11, 12}	{0, 13, 14}	{1, 3, 5}	{1, 4, 7}	{8, 1, 6}
{1, 11, 9}	{1, 10, 13}	{1, 12, 14}	{9, 2, 3}	{2, 4, 6}
{2, 10, 5}	{2, 14, 7}	{8, 2, 12}	{2, 11, 13}	{3, 11, 6}
{3, 12, 7}	{8, 3, 13}	{10, 3, 14}	{4, 5, 13}	{8, 9, 4}
{4, 10, 12}	{11, 4, 14}	{11, 5, 7}	{8, 5, 14}	{9, 12, 5}
{10, 6, 7}	{9, 6, 14}	{12, 13, 6}	{9, 13, 7}	{8, 10, 11}

TABLE VI: Nodes of the Steiner system \mathcal{D}_1

{1, 11, 6}	{1, 2, 5}	{2, 3, 6}	{9, 5, 6}	{3, 11, 5}
{7, 13, 5}	{11, 4, 13}	{2, 12, 7}	{3, 4, 7}	{8, 4, 5}
{8, 11, 7}	{4, 12, 6}	{8, 6, 14}	{12, 5, 14}	{8, 3, 13}
{8, 9, 12}	{9, 10, 13}	{1, 12, 13}	{0, 9, 7}	{9, 2, 11}
{0, 8, 2}	{9, 4, 14}	{10, 11, 14}	{0, 11, 12}	{0, 3, 14}
{8, 1, 10}	{10, 3, 12}	{1, 3, 9}	{0, 10, 5}	{0, 13, 6}
{1, 14, 7}	{2, 4, 10}	{2, 13, 14}	{0, 1, 4}	{10, 6, 7}

TABLE VII: Nodes of the Steiner system \mathcal{D}_2

Let S be a subset of symbols of the design such that no 3-subset of S is contained in a node. By checking all subsets of the symbol set one can observe that the maximum size of S in \mathcal{D}_1 and \mathcal{D}_2 equals 6 ($\{0, 1, 3, 6, 7, 9\}$) and 8 ($\{1, 2, 4, 6, 7, 8, 9, 13\}$) respectively.

This observation results in different file sizes in the codes obtained from the transposes of \mathcal{D}_1 and \mathcal{D}_2 , denoted \mathcal{D}_1^T and \mathcal{D}_2^T respectively. In fact for $k = 7$, the design \mathcal{D}_2^T yields a code which has file size $\mathcal{M}_2 = 28$ which matches the inclusion-exclusion lower bound given by $7 \times 7 - \binom{7}{2}$. However, the design \mathcal{D}_1^T yields a code with file size $\mathcal{M}_1 = 29$ which is strictly larger¹.

We now elaborate on the role of S in the above example. Firstly, note that if \mathcal{D}_i is a Steiner system, then any two storage nodes in \mathcal{D}_i^T intersect in one symbol. Consider the corresponding transposed codes \mathcal{D}_1^T and \mathcal{D}_2^T , where the roles of symbols and nodes is now reversed. As S for \mathcal{D}_2 is of size 8, it implies that we can pick $k = 7$ storage nodes in \mathcal{D}_2^T such that the intersection of any three storage nodes is empty (owing to the definition of S). Thus, upon applying the inclusion-exclusion principle, we obtain the file size to be $7 \times 7 - \binom{7}{2} = 28$.

In contrast, the maximum size of S in \mathcal{D}_1 is 6. Thus, for any set of $k = 7$ storage nodes in \mathcal{D}_1^T there is at least one three-way intersection that is non-empty. Upon exhaustive enumeration, one can realize that the file size in this case is 29 which is strictly higher than 28.

The notion of the set S introduced above can be formalized in terms of a maximal arc in Steiner systems. For Steiner systems that possess a maximal arc, we can therefore determine the file size. In addition, prior results in [32], [33], demonstrate that such maximal arcs exist in a large class of Steiner systems. In the discussion below, we make these arguments in a formal manner.

Definition 12 (s -arc): Let (Ω, V) be a design. A subset $S \subset \Omega$ with $|S| = s$ is called an s -arc if for each node $V_i \in V$ either $|V_i \cap S| = 0$ or $|V_i \cap S| = 2$ holds.

The definition of s -arc implies that any three symbols from S are not contained in any node in V . The largest set S with

¹This example corrects an error in Lemma 11 of [7].

{1, 4, 7, 8}	{0, 2, 8, 9}	{1, 3, 5, 9}	{2, 4, 5, 6}
{0, 3, 6, 7}	{6, 9, 12, 13}	{5, 7, 13, 14}	{6, 8, 10, 14}
{7, 9, 10, 11}	{5, 8, 11, 12}	{2, 3, 11, 14}	{3, 4, 10, 12}
{0, 4, 11, 13}	{0, 1, 12, 14}	{1, 2, 10, 13}	{0, 5, 10, 15}
{1, 6, 11, 15}	{2, 7, 12, 15}	{3, 8, 13, 15}	{4, 9, 14, 15}

TABLE VIII: Nodes of the Steiner system $S(2, 4, 16)$

{1, 4, 12, 13, 15}	{0, 2, 13, 14, 16}	{1, 3, 10, 14, 17}	{2, 4, 10, 11, 18}
{0, 3, 11, 12, 19}	{2, 3, 6, 9, 15}	{3, 4, 5, 7, 16}	{0, 4, 6, 8, 17}
{0, 1, 7, 9, 18}	{1, 2, 5, 8, 19}	{7, 8, 11, 14, 15}	{8, 9, 10, 12, 16}
{5, 9, 11, 13, 17}	{5, 6, 12, 14, 18}	{6, 7, 10, 13, 19}	{15, 16, 17, 18, 19}

TABLE IX: Transposed code obtained from the Steiner system $S(2, 4, 16)$

this property is called a *maximal arc* of the design [34]. It turns out that we can determine the file size for FR codes obtained from transposes of Steiner systems with nontrivial maximal arcs.

For a maximal arc S , consider a symbol $q \in S$. In this case there are $s-1$ pairs of symbols (p, q) such that $p, q \in S$. Since \mathcal{C} is a Steiner system and S is a maximal arc there are $s-1$ distinct nodes in V where each of these pairs occurs. Now, the repetition degree of the system is ρ . Thus, there are $\rho - (s-1)$ nodes which contain the symbol q but no other symbol from S . Based on our assumption, each node in $V_i \in V$ is such that either $|V_i \cap S| = 0$ or $|V_i \cap S| = 2$. Thus, it has to be the case that $s = \rho + 1$.

Lemma 3: Let $\mathcal{C} = (\Omega, V)$ be a FR code derived from a Steiner system $S(2, \alpha, \theta)$ with $\rho = \frac{\theta-1}{\alpha-1}$, such that it has a maximal arc of size $\rho + 1$. Then, the transposed FR code \mathcal{C}^T is such that its code rate is $\frac{k\rho - \binom{k}{2}}{n\alpha}$ for $1 \leq k \leq \rho + 1$.

Proof: In the transposed code \mathcal{C}^T , consider any subset of nodes of size k , where $1 \leq k \leq \rho + 1$. As any two symbols in the original code \mathcal{C} occur in exactly one node of \mathcal{C} it holds that two nodes V_1 and V_2 in \mathcal{C}^T are such that $|V_1 \cap V_2| = 1$. In addition, the storage capacity of the nodes in \mathcal{C}^T is equal to ρ .

Using the inclusion-exclusion principle (cf. Theorem 1), we observe that these nodes cover at least $k\rho - \binom{k}{2}$ symbols in \mathcal{C}^T . Now we pick a set of k nodes in \mathcal{C}^T that correspond to a subset of the maximal arc S in \mathcal{C} . Based on the argument above, it is clear that any two of these nodes intersect in exactly one symbol and any l of the nodes have an empty intersection if $l \geq 3$. It follows that the union of these nodes has exactly $k\rho - \binom{k}{2}$ symbols. The result follows.

Next we provide an explicit example. Let \mathcal{C} be the FR code obtained from a Steiner system $S(2, \alpha = 4, \theta = 16)$.

Example 5 (File size of FR code obtained from the transpose of Steiner System $S(2, \alpha = 4, \theta = 16)$): The nodes in \mathcal{C} are specified in Table VIII and the nodes of the transposed code \mathcal{C}^T are specified in Table IX.

Since the maximal arc should be a set of with cardinality 6, we can choose the symbols greedily and construct the set $S = \{0, 1, 2, 3, 4, 15\}$ as a maximal arc for this Steiner system

According to Lemma 3, the file size for \mathcal{C}^T for $1 \leq k \leq 6$

can be determined by just considering the nodes

$$\{1, 4, 12, 13, 15\}, \{0, 2, 13, 14, 16\}, \{1, 3, 10, 14, 17\}, \\ \{2, 4, 10, 11, 18\}, \{0, 3, 11, 12, 19\}, \text{ and } \{15, 16, 17, 18, 19\}$$

as these correspond to the symbols of S in \mathcal{C} . For these values of k , the file size of the code is $5k - \binom{k}{2}$. Moreover, it is optimal with respect to Singleton bound for $1 \leq k \leq 3$ (cf. Observation 1).

Remark 1 (Steiner Systems with $\alpha = 3, 4$): It is known that several Steiner systems possess maximal arcs. Here we provide the known results for small values of α .

- **(Maximal arcs in Steiner systems with $\alpha = 3$)** By Skolem's construction [35] we have $S(2, 3, \theta)$ for all $\theta \geq 7$ and $\theta \equiv 1, 3 \pmod{6}$. Moreover, for all $\theta \geq 7$ and $\theta \equiv 3, 7 \pmod{12}$ there exists a Steiner system $S(2, 3, \theta)$ with at least one maximal arc [32].
- **(Maximal arcs in Steiner systems with $\alpha = 4$)** It is known [12] that Steiner systems with $\alpha = 4$ exist if and only if $\theta \geq 13$ and

$$\theta \equiv 1, 4 \pmod{12}.$$

Furthermore, if $\rho = \frac{\theta-1}{3}$ is a prime power, then there exists an Steiner system $S(2, 4, \theta)$ with a maximal arc of size $\rho + 1$ [33].

To our best knowledge, there are no other general results about the existence of maximal arcs in Steiner systems with higher values of α .

B. FR codes from resolvable designs

A major drawback of FR codes obtained from Steiner systems is that the repetition degree of the symbols is quite inflexible. In particular, it is not possible to vary the repetition degree and hence the failure resilience of the DSS in an easy way. To address this issue, we now introduce FR codes that are derived from resolvable designs.

A design (Ω, V) is said to be resolvable if we can divide the blocks in V into equal-sized partitions such that (a) each partition contains all the symbols in Ω , and (b) the blocks in a given partition have no symbols in common. Under certain conditions, these designs also allow for β -recoverability. A FR code obtained from such a design is called a resolvable FR code and is naturally resilient to any failure pattern that

ensures that at least one partition is left intact. In the discussion below, we introduce the notion of a net FR code (a subclass of resolvable FR codes) that ensures β -recoverability.

Under this overall framework, we construct several families of net FR codes that allow us to vary the repetition degree in an easy manner. We demonstrate that there exist net FR codes with $\beta > 1$ that cannot be derived by trivial β -expansion. Furthermore, we answer an open question of [7] by demonstrating a FR code that cannot be constructed from Steiner systems. We also provide explicit calculations of the file size for certain ranges of k . The overall structure of this subsection is as follows. We first introduce our construction, show that it results in a net FR code and then calculate its file size.

Definition 13 (Resolvable FR Code): Let $\mathcal{C} = (\Omega, V)$ where $V = \{V_1, \dots, V_n\}$ be a FR code. A subset $P \subset V$ is said to be a parallel class if for $V_i \in P$ and $V_j \in P$ with $i \neq j$ we have $V_i \cap V_j = \emptyset$ and $\cup_{\{j: V_j \in P\}} V_j = \Omega$. A partition of V into r parallel classes is called a resolution. If there exists at least one resolution then the code is called a *resolvable FR code*. For a resolvable FR code, we call two storage nodes parallel if they belong to the same parallel class and non-parallel otherwise. The properties of a resolvable FR code are best illustrated by means of the following example.

Example 6: Consider a DSS with parameters $\alpha = 3, \theta = \alpha^2 = 9, \rho = 2$ and $\beta = 1$. Suppose that we arrange the symbols in $\Omega = \{1, \dots, 9\}$ in a $\alpha \times \alpha$ array A shown below.

$$A = \begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{array}.$$

Let the rows and the columns of A form the nodes in the FR code \mathcal{C} (see Fig. 5), thus $n = 6$. It is evident that there are two parallel classes in \mathcal{C} , $P^r = \{V_1, V_2, V_3\}$ (corresponding to rows) and $P^c = \{V_4, V_5, V_6\}$ (corresponding to columns). As $\rho = 2$, this code can tolerate one failure.

By our construction it is evident that for $V_i \in P^r$ and $V_j \in P^c$, we have $|V_i \cap V_j| = 1$. Using this we can compute the file size \mathcal{M} when $k = 3$, as follows. Let $a + b = 3$ with $a \geq b$. Then, the number of distinct symbols in a set of 3 nodes from \mathcal{C} is

$$3a + (3 - a)(3 - a) = a^2 + 9 - 3a,$$

where a nodes are from P^r and $(3 - a)$ nodes are from P^c . This is minimized when $a = 2$. Thus, $\mathcal{M} = 7$ and $\mathcal{R}_C = \frac{7}{18}$. Note also that the code is optimal with respect to the Singleton bound since $k = \lceil \frac{\mathcal{M}}{\alpha} \rceil$.

If one starts with a resolvable design with many parallel classes, the repetition degree ρ can be varied easily by adding and/or removing parallel classes if needed. We emphasize that the constructions of [7] that are based on Steiner systems largely lack this flexibility as many of them are not resolvable.

In our proposed systems, we require recovery from a node failure by downloading exactly β symbols each from a specified set of d surviving nodes. To address this issue, we consider a subclass of resolvable FR codes called net FR codes where the intersection size of any two nodes from distinct parallel classes is exactly β .

$\{1, 2, 3, 4\}$	$\{5, 6, 7, 8\}$	$\{9, 10, 11, 12\}$	$\{13, 14, 15, 16\}$
$\{1, 5, 9, 13\}$	$\{2, 6, 10, 14\}$	$\{3, 7, 11, 15\}$	$\{4, 8, 12, 16\}$
$\{1, 6, 11, 16\}$	$\{2, 5, 12, 15\}$	$\{3, 8, 9, 14\}$	$\{4, 7, 10, 13\}$
$\{1, 7, 12, 14\}$	$\{2, 8, 11, 13\}$	$\{3, 5, 10, 16\}$	$\{4, 6, 9, 15\}$

TABLE X: A net FR code with parameters $(16, 16, 4, 4)$.

Definition 14 (Net FR Code): Let $\mathcal{C} = (\Omega, V)$ be a resolvable FR code with parameters $(n = ar, \theta = a^2b, \alpha = ab, \rho = r)$ such that any two non-parallel nodes intersect in exactly b symbols. The design determined by \mathcal{C} is called a net [34] and we call \mathcal{C} a *net FR code*.

Examples of net FR codes can be obtained from several combinatorial structures, e.g., grids, affine resolvable designs, Hadamard designs and mutually orthogonal Latin squares (MOLS). We elaborate on these constructions in the subsequent discussion.

Suppose that a net FR code with parameters $(n = ar, \theta = a^2b, \alpha = ab, \rho = r)$ exists. Note that the number of nodes in a parallel class equals $\frac{\theta}{\alpha} = a$. Furthermore, if a given node $V_1 \in V$ fails, this node can be reconstructed by contacting all the nodes in any other intact parallel class and downloading b symbols from each of them. This implies that the code has $d = a, \beta = b$. Next, the code has $\rho = r$ parallel classes and any node can be reconstructed as long as there exists at least one parallel class. Thus, the code is resilient to at least $r - 1$ failures, i.e. $\rho_{res} = r - 1$.

Note that the parameter k can be chosen such that $1 \leq k \leq d$. The code rate \mathcal{R}_C depends on k and needs to be determined. As we shall see determining \mathcal{R}_C can be nontrivial in many cases. Specifically, much of the literature in the area of combinatorial designs focuses on pairwise intersections between the storage nodes, whereas the code rate depends on the minimum size of the intersection of any k storage nodes. Some general results about the code rate of net FR codes can be obtained as discussed in the lemma below. However, a more careful analysis of the algebraic structure of a given construction can allow us to arrive at stronger results.

Lemma 4 (An algorithmic approach for determining the file size of net FR Codes): Let \mathcal{C} be a net FR code with parameters $(n = ar, \theta = a^2, \alpha = a, \rho = r)$, so that $\beta = 1$. Let k be an integer that satisfies $k \leq \rho$ and $\binom{k-1}{2} < a$. Then, the code rate of the system is $\mathcal{R}_C = (\alpha k - \binom{k}{2})/n\alpha$.

Proof: See Appendix.

Example 7: Consider the following FR code obtained from a net with parameters $(n, \theta, \alpha, \rho) = (16, 16, 4, 4)$. The code arises from mutually orthogonal Latin squares (see Section III-B2). This FR code can be specified the nodes presented in Table X. Each row of the table represents a parallel class.

Since any two non-parallel nodes intersect in exactly one point, the code corresponds to a net FR code with $a = 4, b = 1$, and $r = 4$. Thus, $d = 4$ and $\beta = 1$. Suppose that $k = 4$, so that $\binom{k-1}{2} < a$. Our algorithm (cf. Appendix) may choose the following nodes for $k = 4$.

$$L = \{\{1, 2, 3, 4\}, \{1, 5, 9, 13\}, \{2, 5, 12, 15\}, \{4, 6, 9, 15\}\}.$$

So the file size is $\mathcal{M} = 10$. However this code is not optimal with respect to Singleton bound. However, observe that the

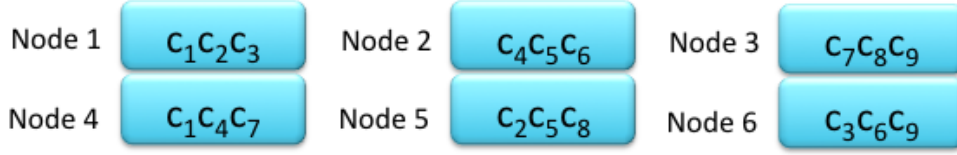


Fig. 5: A DSS specified with $(n = 6, k = 3, d = 3, \alpha = 3)$. Note that the nodes numbered 1,2,3 and 4,5,6 form parallel classes.

code formed by deleting a parallel class has parameters $(n = 12, \theta = 16, \alpha = 4, \rho = 3)$. In this code any three nodes cover at least 9 symbols. Thus setting $k = \lceil \frac{9}{4} \rceil = 3$ (cf. Observation 1) results in a code that meets the Singleton bound.

Note that while Lemma 4 applies to all net FR codes with $\beta = 1$, the requirement that the storage capacity $\alpha = a \geq \binom{k-1}{2}$ is quite restrictive. For certain net FR codes that have a tractable algebraic and/or geometric characterization we can perform a more careful analysis and we now turn our attention to them. Our first example is a net where the file size \mathcal{M} is strictly larger than $k\alpha - \beta \binom{k}{2}$.

1) *Affine Resolvable FR code*: Affine resolvable designs are a class of resolvable designs where the intersection between two nodes in different parallel classes can be computed exactly. These can be derived from affine geometries that can be intuitively understood as follows. The set of points corresponds to all elements of \mathbb{F}_q^n , the vector space of dimension n over a finite field of size q , \mathbb{F}_q . Thus, the number of points is q^n . The blocks correspond to the solutions of certain sets of linear equations over the vector space. For the sake of simplicity, let us consider just one equation, e.g., $x_1 = a$ for $a \in \mathbb{F}_q$. For each $a \in \mathbb{F}_q$ the solution set is of size q^{n-1} . Each such solution set corresponds to a block in the design. Furthermore, these solution sets partition \mathbb{F}_q^n . In a similar manner, one can consider other sets of linear equations of the form $\sum_{i=1}^n b_i x_i = a$ where $b_i \in \mathbb{F}_q$ whose solution sets also partition \mathbb{F}_q^n . Furthermore any two linear independent linear equations will have a solution set of size q^{n-2} , i.e., the intersection between two such blocks will be exactly q^{n-2} .

The resultant block design is a resolvable design [12]. In the discussion below, we present a formal presentation of this idea. We also analyze the file size of the obtained system under the condition that the equations are chosen in a specific manner and for an appropriate range of k .

Let q be a prime power, $m \geq 2$ and $\Omega = \mathbb{F}_q^m$. Let $1 \leq \delta \leq m-1$. We treat Ω as an m -dimensional vector space over \mathbb{F}_q . A δ -flat is the solution set to a system of $m-\delta$ independent linear equations that can be homogeneous or non-homogeneous. The set Ω and the set of all δ -flats of Ω comprise the m -dimensional affine geometry over \mathbb{F}_q , denoted by $AG_m(q)$. It turns out that one can generate a large class of resolvable designs by considering $AG_m(q)$. Let $\begin{bmatrix} m \\ \delta \end{bmatrix}_q$ denote the Gaussian coefficient, so that

$$\begin{bmatrix} m \\ \delta \end{bmatrix}_q = \begin{cases} \frac{(q^m-1)(q^{m-1}-1)\dots(q^{m-\delta+1}-1)}{(q^\delta-1)(q^{\delta-1}-1)\dots(q-1)} & \text{if } \delta \neq 0, \\ 1 & \text{if } \delta = 0. \end{cases}$$

Theorem 2 (Affine Resolvable Designs): [12] Let V denote the set of all δ -flats in $AG_m(q)$. Then $\Omega = \mathbb{F}_q^m$ and V form

$R_1 = \{000, 001, 002, 010, 020, 011, 012, 021, 022\}$
$R_2 = \{000, 001, 002, 100, 200, 101, 102, 201, 202\}$
$R_3 = \{000, 001, 002, 110, 220, 111, 112, 221, 222\}$
$R_4 = \{000, 001, 002, 120, 210, 121, 122, 211, 212\}$
$R_5 = \{000, 010, 020, 100, 200, 110, 120, 210, 220\}$
$R_6 = \{000, 010, 020, 101, 202, 111, 121, 212, 222\}$
$R_7 = \{000, 010, 020, 102, 201, 112, 122, 211, 221\}$
$R_8 = \{000, 011, 022, 100, 200, 111, 122, 211, 222\}$
$R_9 = \{000, 011, 022, 101, 202, 112, 120, 210, 221\}$
$R_{10} = \{000, 011, 022, 102, 201, 110, 121, 212, 220\}$
$R_{11} = \{000, 012, 021, 100, 200, 112, 121, 212, 221\}$
$R_{12} = \{000, 012, 021, 101, 202, 110, 122, 211, 220\}$
$R_{13} = \{000, 012, 021, 102, 201, 111, 120, 210, 222\}$

TABLE XI: Representatives of parallel classes of the FR code with parameters $(n = 39, \theta = 27, \alpha = 9, \beta = 3, d = 3, \rho = 13)$.

a resolvable BIBD with $(\theta = q^m, \rho, \alpha = q^\delta, \lambda)$ -BIBD with $n = q^{m-\delta} \begin{bmatrix} m \\ \delta \end{bmatrix}_q, \rho = \begin{bmatrix} m \\ \delta \end{bmatrix}_q$ and $\lambda = \begin{bmatrix} m-1 \\ \delta-1 \end{bmatrix}_q$.

The case of $m = 2, \delta = 1$ corresponds to affine planes. When $\delta = m-1$ we obtain an affine resolvable BIBD with $n = \theta + \rho - 1$. In this case the DSS is specified by the parameters $\theta = q^m, \alpha = q^{m-1}, \rho = \frac{q^m-1}{q-1}$ and $n = q\rho$. The design can be obtained by means of the following algorithm.

- (i) Let $\Omega = \{(x_1, x_2, \dots, x_m) : x_i \in \mathbb{F}_q \text{ for } i = 1, 2, \dots, m\}$ be the symbol set.
- (ii) Find $\rho, (m-1)$ -dimensional subspaces of \mathbb{F}_q^m such that each of them contains the symbol $(0, 0, \dots, 0) \in \mathbb{F}_q^m$. Note that these subspaces of \mathbb{F}_q^m are the solutions to a single homogeneous linear equation over \mathbb{F}_q in q variables. These ρ subspaces are representatives of the ρ different parallel classes.
- (iii) Construct each parallel class by considering the additive cosets of its representative. Let R_1 be a $(m-1)$ -dimensional subspace corresponding to a given homogeneous equation. Let $U = \{0, u_1, \dots, u_{q-1}\}$ be the full set of coset representatives of R_1 . The rest of the blocks can be obtained by the cosets $R_1^i = u_i + R_1$. Note that each of these cosets corresponds to a nonhomogeneous equation.

Example 8 (An example of an Affine Resolvable Design): [12] Let $q = 3$ and $m = 3$. The set of symbols is $\Omega = \mathbb{F}_3^3$ and there are 39 blocks which can be partitioned into 13 parallel classes. The representatives of the 13 parallel classes are specified in the Table XI, where the vector $[x_1 \ x_2 \ x_3]$ is simply written as $x_1 x_2 x_3$. The other blocks are additive cosets of these 13 representatives. For example, the first parallel class

consists of the following blocks.

$$\begin{aligned} B_1 &= \{000, 001, 002, 010, 020, 011, 012, 021, 022\}, \\ B_2 &= \{100, 101, 102, 110, 120, 111, 112, 121, 122\}, \text{ and} \\ B_3 &= \{200, 201, 202, 210, 220, 211, 212, 221, 222\}. \end{aligned}$$

Here the blocks B_1 , B_2 and B_3 correspond to equations $x_1 = 0$, $x_1 = 1$ and $x_1 = 2$ respectively.

The overlap between blocks from different parallel classes in the case of affine resolvable designs is known from the following result.

Lemma 5: [12] Any two blocks from different parallel classes of an affine resolvable $(\theta, \rho, \alpha, \lambda)$ -BIBD intersect in exactly α^2/θ symbols.

Using the above facts, we can conclude that an affine resolvable BIBD is an instantiation of a net FR code with parameters $(n = q^{\frac{q^m-1}{q-1}}, \theta = q^m, \alpha = q^{m-1}, \rho = \frac{q^m-1}{q-1})$ and $d = q, \beta = q^{m-2}$. Of course, the repetition degree can be varied by only retaining as many parallel classes as needed.

Remark 2 (Affine Resolvable FR Codes cannot be obtained by trivial β -expansion): It is important to note that the affine resolvable FR codes are an example of a FR code family with $\beta > 1$ that cannot be obtained by replicating the symbols of a smaller code. To show this we will simply use Observation 2. Specifically, consider $m \geq 2q + 1$ and $q \geq 3$. In this case the affine resolvable FR code will have parameters $\theta = q^m, \alpha = q^{m-1}, \rho = \frac{q^m-1}{q-1}, n = q\rho$ and $\beta = q^{m-2}$. If it could be generated from a smaller code simply by replication, this would imply that the smaller code had a storage capacity of q and q^2 total symbols. This means it has at most $\binom{\theta/\beta}{\alpha/\beta} = \binom{q^2}{q} \leq \left(\frac{q^2 e}{q}\right)^q = (eq)^q \leq q^{2q}$ distinct storage nodes. However, in the affine resolvable FR code we have $n = q^{\frac{q^m-1}{q-1}} \geq q^{\frac{q^{2q+1}-1}{q-1}}$ which can be verified to be strictly larger than q^{2q} .

We can determine the file size of a code \mathcal{C} obtained from some specific affine resolvable designs, for certain ranges of k . We consider two scenarios depending on the relationship between q and m .

- **(Case 1: $q > m$)**

We choose the code \mathcal{C} such that it has $r \geq m$ parallel classes such that the i -th parallel class of \mathcal{C} corresponds to the homogeneous equation $x_1 + \alpha_i x_2 + \alpha_i^2 x_3 + \dots + \alpha_i^{m-1} x_m = 0$, where $\alpha_i, i = 1, \dots, r$ are all non-zero and distinct. Note that the distinctness requirement also enforces that $q > r$. The equations obtained in this manner are such that any m equations are linearly independent [36].

For this code we analyze the file size for a fixed $k \leq m$. For a given set of k blocks, denoted $A_i, i = 1, \dots, k$, it is possible that multiple blocks from the same parallel class are chosen; suppose that these blocks come from l distinct parallel classes, numbered without loss of generality as $1, \dots, l$. Let z_i denote the number of blocks from the i -th parallel class, so that

$$z_1 + z_2 + \dots + z_l = k.$$

If we pick k_1 blocks each from a different parallel class, we can immediately conclude that the total number of symbols covered is q^{m-k_1} , as the parallel classes correspond to linearly independent equations. Using this fact and the inclusion-exclusion principle, we have

$$\begin{aligned} |\cup_{i=1}^k A_i| &= \sum_{i_1=1}^l z_{i_1} q^{m-1} - \sum_{i_1 < i_2} z_{i_1} z_{i_2} q^{m-2} \\ &+ \sum_{i_1 < i_2 < i_3} z_{i_1} z_{i_2} z_{i_3} q^{m-3} + \dots + (-1)^l z_1 z_2 \dots z_l q^{m-l}. \end{aligned}$$

Upon inspection, it is clear that

$$\begin{aligned} q^m \left(1 - \prod_{i=1}^l \left(1 - \frac{z_i}{q} \right) \right) &= \sum_{i_1=1}^l z_{i_1} q^{m-1} - \sum_{i_1 < i_2} z_{i_1} z_{i_2} q^{m-2} \\ &+ \sum_{i_1 < i_2 < i_3} z_{i_1} z_{i_2} z_{i_3} q^{m-3} + (-1)^l z_1 z_2 \dots z_l q^{m-l}. \end{aligned} \quad (5)$$

Thus, we need to analyze the minimum value of the LHS of equation (5) (over the possibilities for $z_i, i = 1, \dots, l$) to determine the file size. Using the AM-GM inequality, we obtain

$$\begin{aligned} \frac{1}{l} \sum_{i=1}^l \left(1 - \frac{z_i}{q} \right) &= 1 - \frac{k}{lq} \geq \left[\prod_{i=1}^l \left(1 - \frac{z_i}{q} \right) \right]^{\frac{1}{l}} \\ \implies \left[1 - \frac{k}{lq} \right]^l &\geq \prod_{i=1}^l \left(1 - \frac{z_i}{q} \right). \end{aligned}$$

Equality holds in the above equation when all the z_i terms are equal. In addition, we show below that the function

$$h(l) = \left[1 - \frac{k}{lq} \right]^l$$

takes its maximum value over the set $l = 1, \dots, k$ when $l = k$. To see this, let $0 < \chi = \frac{k}{q} < 1$, and consider $\log h(l) = l \log(1 - \frac{\chi}{l})$. Now,

$$\frac{d}{dl} \log h(l) = \log(1 - \frac{\chi}{l}) + \frac{\frac{\chi}{l}}{1 - \frac{\chi}{l}}.$$

Let $\chi_1 = \frac{\chi}{l}$ and let us study the function $h_1(\chi_1) = \log(1 - \chi_1) + \frac{\chi_1}{1 - \chi_1}$. Clearly $h_1(0) = 0$. The derivative of $h_1(\chi_1)$ is non-negative for $0 < \chi_1 < 1$, since it equals $\frac{\chi_1}{(1 - \chi_1)^2}$. This implies that $h_1(\chi_1) \geq 0$ for $0 < \chi_1 < 1$ and therefore $h'(l) \geq 0$ in the range $l = 1, \dots, k$, i.e., it is an increasing function in this range. This implies that the maximum value of $h(l)$ in the range $l = 1, \dots, k$ is obtained when $l = k$ and $z_i = 1$ for all i .

We conclude that the minimum value of the LHS of equation (5) is obtained when $k = l$ and $z_i = 1, i = 1, \dots, k$ and that the file size is $q^m \left(1 - \left(1 - \frac{1}{q} \right)^k \right)$.

- **(Case 2: $q \leq m$)**

In this case we choose the code \mathcal{C} so that it has $r \leq m$ parallel classes. The chosen parallel classes are such that they belong to linearly independent equations. Once again, we can analyze the file size when $k \leq m$. Suppose that we choose l parallel classes and let z_i denote the number of blocks chosen from the i -th parallel class. Note

that in this case $l \geq \lceil \frac{k}{q} \rceil$ and $z_i \leq q$ for all $i = 1, \dots, l$. Proceeding as in Case 1, we can argue that the function

$$h(l) = \left[1 - \frac{k}{lq}\right]^l$$

attains its maximum when $l = k$ and $z_i = 1$ for all $i = 1, \dots, k$. Thus, in this case as well the maximum file size is given by $q^m \left(1 - \left(1 - \frac{1}{q}\right)^k\right)$.

2) *Resolvable FR codes from Grids, Hadamard designs and MOLS*: Note that affine resolvable codes have β which is a prime power. We now construct families of net FR codes where $\beta = 1$. Overall, the idea here is to relate the existence of these codes to combinatorial structures such as grids (two-dimensional arrays), Hadamard designs and mutually orthogonal Latin squares. While these combinatorial structures have been studied in their own right, their usage in constructing FR codes is new. In particular, our construction from MOLS demonstrates an instance of a FR code that cannot be derived from Steiner systems (answering an open question in [7]).

An $a \times a$ grid is a FR code that is obtained as follows.

- Let $\Omega = \{0, \dots, a^2 - 1\}$. Create an 2D-array A whose (i, j) -th entry is $a \times i + j$, where $0 \leq i, j \leq a - 1$.
- Each column and each row of A determines a storage node.

It is clear that the FR code so obtained is resolvable. Specifically, the set of columns and the set of rows form a resolution. The parameters are $(n = 2a, \theta = a^2, \alpha = a, \rho = 2)$. Note that $\beta = 1$ as any row and any column intersect in exactly one symbol. Thus, the code so obtained is also a net FR code.

Lemma 6 (File size of grid FR Codes): Let \mathcal{C} be a net FR code obtained from an $a \times a$ grid. If k is even, the file size \mathcal{M} of \mathcal{C} is $ka - k^2/4$ and if k is odd, it is $ka - (k^2 - 1)/4$.

Proof: Assume that we choose s nodes from the parallel class corresponding to the rows and t nodes from the parallel class corresponding to the columns such that $s + t = k$. Note that $k \leq d = a$. It is evident that any three nodes have an empty intersection. Thus, applying the inclusion-exclusion principle, we conclude that any k nodes cover exactly $\alpha k - st$ symbols. Next, note that $\alpha k - st = ak - ks + s^2 = (s - k/2)^2 + ka - k^2/4$ which takes the minimum value $ka - k^2/4 + \min((k/2 - \lceil k/2 \rceil)^2, (k/2 - \lfloor k/2 \rfloor)^2)$, i.e., it equals $ka - k^2/4$ when k is even and $ka - (k^2 - 1)/4$ when k is odd.

The following corollary can be obtained by examining conditions under which $k = \lceil \frac{\mathcal{M}}{\alpha} \rceil$.

Corollary 1:

- Let $k = 2u$ and $u^2 < a$. Then the FR code obtained from $a \times a$ grid is optimal with respect to the Singleton bound.
- Let $k = 2u + 1$ and $u(u + 1) < a$. Then the FR code obtained from $a \times a$ grid is optimal with respect to the Singleton bound.

A second construction of affine resolvable designs can be obtained from Hadamard matrices or equivalently difference sets as discussed below. Consider an algebraic group G of order θ and $D \subseteq G$ such that $|D| = \alpha$, with the property that every nonidentity element of G can be expressed as a

$\{\infty, 1, 2, 4\}$	$\{0, 3, 5, 6\}$
$\{\infty, 2, 3, 5\}$	$\{1, 4, 6, 0\}$
$\{\infty, 3, 4, 6\}$	$\{2, 5, 0, 1\}$
$\{\infty, 4, 5, 0\}$	$\{3, 6, 1, 2\}$
$\{\infty, 5, 6, 1\}$	$\{4, 0, 2, 3\}$
$\{\infty, 6, 0, 2\}$	$\{5, 1, 3, 4\}$
$\{\infty, 0, 1, 3\}$	$\{6, 2, 4, 5\}$

TABLE XII: Hadamard design obtained from the $(7, 3, 1)$ -difference set in $\Omega = \mathbb{F}_7$.

difference $d_1 - d_2$ of elements of D in exactly λ ways. We refer to D as a $(\theta, \alpha, \lambda)$ -difference set.

Lemma 7 (Quadratic Residue Difference Set): [12] Let $q = 4a - 1 \geq 7$ be an odd prime power and $G = \mathbb{F}_q$. Let $D = \{z^2 : z \in \mathbb{F}_q, z \neq 0\}$ be the set of quadratic residues. Then D is a $(4a - 1, 2a - 1, a - 1)$ -difference set in $(\mathbb{F}_q, +)$, where $+$ denotes the additive operation over \mathbb{F}_q .

For any $g \in G$, we define the *translate* of D by $g + D = \{g + d : d \in D\}$, and define the *development* of D by $\text{Dev}(D) = \{g + D : g \in G\}$. If D is a $(\theta, \alpha, \lambda)$ -difference set in G , then $(G, \text{Dev}(D))$ is a $(\theta, \rho, \alpha, \lambda)$ -BIBD [12].

Let (Ω, V) be the $(4a - 1, 2a - 1, 2a - 1, a - 1)$ -BIBD constructed by using a quadratic residue difference set. Let $\infty \notin \Omega$, and define for $V' = \{B \cup \{\infty\} : B \in V\}$. Then it can be shown that $(\Omega \cup \{\infty\}, V' \cup \{\Omega - B : B \in V\})$ is an affine resolvable $(4a, 4a - 1, 2a, 2a - 1)$ -BIBD. Using the equations (3) and (4) this corresponds to a net FR code with parameters $\theta = 4a, \alpha = 2a, \beta = a, d = 2, \rho = 4a - 1$ and $n = 8a - 2$ (see [12], Chapter 5).

Example 9: $D = \{1, 2, 4\}$ is a $(7, 3, 1)$ -difference set in $\Omega = \mathbb{F}_7$. We can construct the Fano plane by using the difference set D which is a $(7, 3, 3, 1)$ -BIBD. By applying the above construction we can construct a FR code with parameters $\theta = 8, n = 14, \alpha = 4, \rho = 7$. Corresponding storage nodes are presented in Table XII where each row of the table represents a parallel class.

For this class of codes, d is always 2. However, they offer more flexibility in the choice of β ; unlike affine geometry based codes, we do not require β to be a prime power.

Remark 3 (FR Codes derived from Hadamard Designs cannot be obtained by trivial β -expansion with $\beta = a$): In addition, they provide another example of a family of FR codes that cannot be obtained by trivial β -expansion with $\beta = a$. To show this, we use Observation 2. Suppose that such a code could be obtained by trivial β -expansion with $\beta = a$, then the original code would correspond to a FR code with 4 symbols and storage capacity of 2. In this case, there can be at most $\binom{4}{2} = 6$ nodes. In contrast, the code obtained from the Hadamard design has $8a - 2 > 6$ nodes (as $a \geq 2$).

Since any two non-parallel nodes share a symbols in common, any $k = 2$ nodes cover at least $3a$ symbols where $\alpha = 2a$. Moreover, $k = 2 = \lceil \frac{3a}{2a} \rceil$. Hence the code is optimal with respect to Singleton bound for $k = 2$.

We now discuss another construction of net FR codes that can be obtained from MOLS.

Definition 15 (Latin Square): A Latin square of order a with entries from a set Ω with $|\Omega| = a$ is an $a \times a$ array L

in which every cell contains an element of Ω such that every row of L is a permutation of Ω and every column of L is a permutation of Ω .

Definition 16 (Orthogonal Latin Squares): Suppose that L_1 and L_2 are Latin squares of order a with entries from Ω_1 and Ω_2 respectively (where $|\Omega_1| = |\Omega_2|$). We say that L_1 and L_2 are orthogonal Latin squares if for every $x \in \Omega_1$ and for every $y \in \Omega_2$ there is a unique cell (i, j) such that $L_1(i, j) = x$ and $L_2(i, j) = y$.

Equivalently, one can consider the superposition of L_1 and L_2 in which each cell (i, j) is occupied by the pair $(L_1(i, j), L_2(i, j))$. Then, L_1 and L_2 are orthogonal if and only if the resultant array has every value in $\Omega_1 \times \Omega_2$. A set of r Latin squares L_1, \dots, L_r of order a are said to be *mutually orthogonal* if L_i and L_j are orthogonal for all $1 \leq i < j \leq r$.

We now demonstrate a procedure of constructing net FR codes from MOLS [37]. Let $\Omega = \{1, 2, \dots, a^2\}$, and let L_1, L_2, \dots, L_{r-2} be a set of $r-2$ MOLS of order a ($r-2 \leq a-1$).

- Arrange the elements of Ω in a $a \times a$ array A . Each row and each column of A corresponds to a storage node (this gives us $2a$ nodes).
- Note that L_i takes values in $\{1, \dots, a\}$. Within L_i identify the set of (i, j) pairs where a given value $z \in \{1, \dots, a\}$ appears. Create a storage node by including the entries of A corresponding to the identified (i, j) pairs.
- Repeat this for each L_i and all $z \in \{1, \dots, a\}$. This creates another $(r-2)a$ storage nodes.

Thus, a total of ra storage nodes of size a can be obtained. Of course, one can choose fewer storage nodes if so desired.

Example 10: Let $a = 4$, and $r = 2$. Then, we have the following construction.

$$A = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{array}, \quad L_1 = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{array} \text{ and } L_2 = \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \\ 2 & 1 & 4 & 3 \end{array}.$$

We have the cells $(L_1(i, j), L_2(i, j))$ for $i, j = 1, 2, 3, 4$ in a matrix form as follows:

$$\begin{array}{cccc} (1, 1) & (2, 2) & (3, 3) & (4, 4) \\ (2, 3) & (1, 4) & (4, 1) & (3, 2) \\ (3, 4) & (4, 3) & (1, 2) & (2, 1) \\ (4, 2) & (3, 1) & (2, 4) & (1, 3) \end{array}.$$

As we can see from this matrix, all possible cells are covered by the cells $(L_1(i, j), L_2(i, j))$. Thus L_1 and L_2 are orthogonal. We have the parallel classes and corresponding storage nodes illustrated in Example 7.

Note that in describing the above construction we assumed the existence of $r-2$ MOLS. We now discuss the issue of the existence of such structures. If p is a prime number, m is a positive integer, and $N = p^m$ then we can construct $N-1$ mutually orthogonal Latin squares as described below.

- Define $L_a : \mathbb{F}_N \times \mathbb{F}_N \rightarrow \mathbb{F}_N$, by $(r, c) \mapsto ar + c$ (where the addition is over \mathbb{F}_N) for all $a \in \mathbb{F}_N \setminus \{0\}$. Then, L_a is a Latin square since for a given row r (or column c) the column (or row) location of an element s is uniquely specified.
- For any $a, b \in \mathbb{F}_N \setminus \{0\}$, L_a and L_b are orthogonal since for given ordered pair (s, t) the system $ar + c = s, br + c = t$, determine $r = (a-b)^{-1}(s-t)$ and $c = s - ar$ uniquely.

Example 11: Let $N=3$. Then $\mathbb{F}_3 = \{0, 1, 2\}$, $L_1 : x + y$ and $L_2 : 2x + y$. The two orthogonal Latin squares of order 3 constructed by the above method are

$$L_1 = \begin{array}{ccc} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{array}, \text{ and } L_2 = \begin{array}{ccc} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{array}$$

It turns out that in general, the construction described above produces a net FR code. The parameters are discussed in the following discussion.

Lemma 8: The construction procedure described above produces a net FR code with $\theta = a^2, n = ra, d = \alpha = a, \rho = r$ where non-parallel nodes intersect in exactly one point.

Proof: It is clear from the construction that $\theta = a^2$ and $n = ra$. Each storage node has a symbols so that $\alpha = a$. We need to show that the code is resolvable. Towards this end, note that it is evident that we obtain a parallel class by considering the nodes corresponding to the rows of A (a similar argument holds for the columns of A). Next, the nodes obtained by considering Latin square L_i also form a parallel class, since the set of elements obtained by considering the (i, j) pairs corresponding to $z_1 \in \{1, \dots, a\}$ are distinct from those corresponding to $z_2 \in \{1, \dots, a\}$, if $z_1 \neq z_2$. As we have r parallel classes, we obtain $\rho = r$. Next, consider the overlap between any two storage nodes belonging to different parallel classes. As L_i and L_j are orthogonal, any entry $(k, l) \in [a] \times [a]$ appears exactly once in the superposition of L_i and L_j , which implies that the overlap between storage nodes from different parallel classes corresponding to the L_i 's is exactly one element. Similarly, a block from a parallel class corresponding to L_i has exactly one overlap with the blocks corresponding to the rows and columns of A .

Remark 4 (There are FR Codes which can be obtained from MOLS but not from Steiner Systems): In general, the construction of orthogonal Latin squares is somewhat involved. However, the celebrated results of [11], demonstrate the construction of two orthogonal Latin squares for all orders $N \neq 2, 6$. This immediately allows us to construct net FR codes with the following parameters $n = 4a, \theta = a^2, d = \alpha = a, \beta = 1$, and $\rho = 4$ for any $a \neq 2, 6$. By applying Lemma 4 we can get the file size $\mathcal{M} = 4a - 6$ for $k = 4$ for $a > 6$ and it is optimal with respect to Singleton bound (cf. Observation 1).

This construction allows us to design some FR codes whose parameters cannot be obtained from Steiner systems. For instance, Let $\alpha = 10$ and $\theta = 100$. Then to construct a FR code we need use the Steiner system $S(2, 10, 100)$ which does not exist [38]. However the above construction with two orthogonal Latin squares of order 10 provides us a net FR

code with $\alpha = 10$ and $\theta = 100$.

Lemma 9 (File size of FR Codes obtained from MOLS):

Let p be a prime and m be a positive integer, so that there exist $p^m - 1$ MOLS of order p^m . Consider a subset of these $p^m - 1$ MOLS of size r and let \mathcal{C} be a net FR code constructed from them. Then for any $k \leq r$, the code rate $R_{\mathcal{C}} = (k(p^m) - \binom{k}{2})/np^m$.

Proof: Let η be a primitive element of \mathbb{F}_{p^m} . From the construction of the r MOLS, we can associate a set of non-zero field elements $\{\eta^{\alpha_1}, \dots, \eta^{\alpha_r}\}$ so that the i -th Latin square is generated by the corresponding η^{α_i} , where α_i 's are distinct. In the discussion below we demonstrate the existence of r storage nodes that cover exactly $rp^m - \binom{r}{2}$ symbols. The argument will also show the required result for any $k < r$. From the inclusion-exclusion principle it is evident that any r nodes cover at least $rp^m - \binom{r}{2}$ symbols. For demonstrating a set of nodes that cover exactly this number we first pick the storage nodes from different parallel classes and demonstrate that the intersection of any three nodes from this set is empty.

Towards this end in the i -th MOLS, consider the storage node determined by the equation $\eta^{\alpha_i}x + y = \eta^{2\alpha_i}$. This specifies the set of nodes that we will be considering. Three nodes intersect in some symbol if the following system of equations has a solution.

$$\eta^{\alpha_i}x + y = \eta^{2\alpha_i} \quad (6)$$

$$\eta^{\alpha_j}x + y = \eta^{2\alpha_j} \quad (7)$$

$$\eta^{\alpha_k}x + y = \eta^{2\alpha_k} \quad (8)$$

Note that any two equations from the set above are linearly independent and have exactly one solution. Thus, if the above system has a solution, then there exist $\mu \neq 0$ and $\lambda \neq 0$ such that

$$\lambda\eta^{\alpha_i} + \mu\eta^{\alpha_j} = \eta^{\alpha_k}$$

$$\lambda + \mu = 1$$

$$\lambda\eta^{2\alpha_i} + \mu\eta^{2\alpha_j} = \eta^{2\alpha_k}$$

Next, we note that it cannot be the case that $\eta^{2\alpha_i} = \eta^{2\alpha_j} = \eta^{2\alpha_k}$. To see this note that there are no zero divisors in a finite field so $z_1^2 = z_2^2$ implies $z_1 = z_2$ or $z_1 = -z_2$. Thus, we can conclude that

$$\lambda = \frac{\eta^{\alpha_k} - \eta^{\alpha_j}}{\eta^{\alpha_i} - \eta^{\alpha_j}} = \frac{\eta^{2\alpha_k} - \eta^{2\alpha_j}}{\eta^{2\alpha_i} - \eta^{2\alpha_j}}.$$

However

$$\frac{\eta^{2\alpha_k} - \eta^{2\alpha_j}}{\eta^{2\alpha_i} - \eta^{2\alpha_j}} = \frac{(\eta^{\alpha_k} - \eta^{\alpha_j})(\eta^{\alpha_k} + \eta^{\alpha_j})}{(\eta^{\alpha_i} - \eta^{\alpha_j})(\eta^{\alpha_i} + \eta^{\alpha_j})}$$

and this implies $\eta^{\alpha_i} = \eta^{\alpha_k}$ which is a contradiction. Thus, a solution to the system of equations in (6) - (8) does not exist. The result follows.

Remark 5: The existence of $p^m - 1$ MOLS implies the existence of an affine plane of order p [12]. Thus choosing $r = p^m - 1 = k$, we can obtain the corresponding file sizes for affine planes. Codes constructed from affine planes were also considered in [7] under Steiner systems.

Example 12: A FR code obtained from affine plane of order 3 is depicted in Fig. 6. This code can be obtained by following the construction outlined above with $p = 3$ and $m = 1$. It can be observed that this code is optimal with respect to the Singleton bound when $k = 2$. (cf. Observation 1).

C. Discussion of code parameters achieved by the proposed constructions

In this subsection, we summarize the range of DSS parameters that our constructions can achieve. Note that there are certain parameter restrictions that any FR code has to satisfy. We list these below. To avoid trivialities, we assume there are no repeated storage nodes in the system.

$n\alpha = \theta\rho$, by counting the number of ones in the incidence matrix,

$n \leq \binom{\theta}{\alpha}$, as the nodes are α -sized subsets of the symbols,

$2 \leq \alpha \leq \theta - 1$, as the storage capacity can be at most $\theta - 1$,

$1 < k \leq d = \frac{\alpha}{\beta}$, as the nodes need to be β -recoverable,

and

$$\alpha + 1 \leq \mathcal{M} \leq k\alpha.$$

If $\beta = 1$, the result of [29] shows that the conditions are also sufficient for the existence of a FR code; however [29] does not discuss the file size of such a code. It is evident that specific construction technique imposes additional restrictions. For instance, if the FR code is obtained from a resolvable design, then $\frac{\theta}{\alpha}$ needs to be an integer as it is the number of nodes in a parallel class. In Tables I – V (cf. section II), we summarize the parameters (and the corresponding restrictions that apply) of the different constructions proposed above.

We emphasize that any FR code is equivalent to a biregular bipartite graph (cf. Definition 8) and the file size for a given value of k is closely related to the expansion properties of k -sized subsets of the storage nodes. It is well recognized that determining the expansion of an arbitrary bipartite graph is a computationally hard problem. In particular, precise numbers are known only for certain families of graphs. High probability results for expansion are known; however, such results are asymptotic in nature and do not provide deterministic constructions. Parameters such the file size can only be found by inspection of the randomly constructed graph. Furthermore, it is not clear whether the β -recoverability property can be shown for these codes. For these reasons, it is very hard to fully characterize the range of achievable parameters for FR codes (other than the necessary constraints presented above).

Reference [30] presents results on the file size of resolvable FR codes that we have considered above. However, we emphasize that Theorem 19 in [30] does not apply in our situation. For instance, consider the construction of FR codes from MOLS presented above and Lemma 9. Suppose that we choose $r = p^m - 1$ and $k = p^m - 2$. In this case it can be verified that for large p , the result of Theorem 19 in [30] does not apply. Furthermore, our affine resolvable design based construction has $\beta > 1$ and the results of [30] do not apply here.

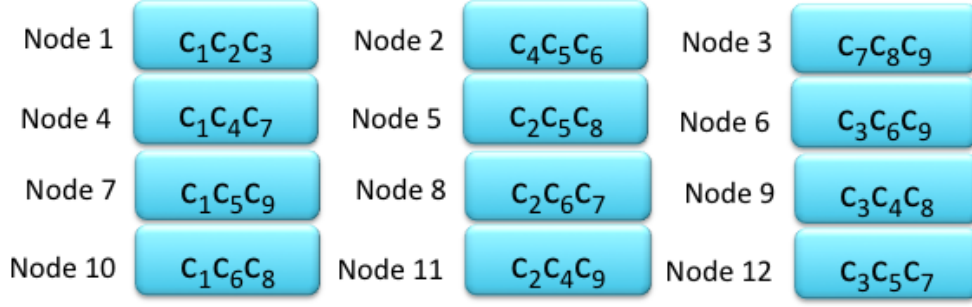


Fig. 6: FR code derived from an affine plane of order 3.

On a different note, it can also be argued that one can simply treat the FR codes discussed in this section as local codes, by choosing a value of k that is strictly larger than d (note that k is under our control as a system designer). However, we will now argue that this will result in significantly suboptimal codes with respect to the minimum distance bound in Lemma 2. Suppose for instance that we consider a net FR code with parameters $(n = ar, \theta = a^2b, \alpha = ab, \rho = r)$ with $\beta = b$ and $d = a$. Note that there are r parallel classes in the code. The bound in eq. (2), reduces to the Singleton bound as $d\alpha = a^2b = \theta$, so that $\lceil \frac{\mathcal{M}}{d\alpha} \rceil = \lceil \frac{\mathcal{M}}{\theta} \rceil = 1$. Thus, while increasing the value of k above d makes the code local, it will be far from the achieving the local code minimum distance bound in Lemma 2. As a concrete example, consider a grid code (an instantiation of the net FR code) with $(n = 20, \theta = 100, \alpha = 10, \rho = 2)$. In this case $d = 10$ and if $k = 6$, the code is optimal with respect to the Singleton bound as $d_{\min} = 20 - \lceil \frac{51}{10} \rceil + 1 = 15$. However if choose $k = 11$, so that it becomes a local code, the corresponding file size is $\mathcal{M} = 80$, so that the minimum distance bound is $20 - \lceil \frac{80}{10} \rceil + 1 = 13$. However, this code can only recover from at most 9 node failures and not 12. Thus, such a code is a suboptimal local regenerating code. As all the resolvable codes presented in this section are instances of net FR codes, similar statements apply to all these constructions.

IV. SOME CHARACTERISTICS OF FR CODES OBTAINED FROM KRONECKER PRODUCTS

The resolvable FR codes derived from affine resolvable designs and Hadamard designs are families of FR codes that have $\beta > 1$ and in many cases cannot be obtained via trivial β -expansion. In this section, we present the Kronecker product as a technique for obtaining new codes that have $\beta > 1$. In essence, we demonstrate the following result. Suppose that we start with a base FR code with storage capacity α where the pairwise intersection between storage nodes is at most one symbol and is such that its file size equals the inclusion-exclusion lower bound in eq. (2). If we consider the Kronecker product of the code with itself, we get a new FR code, where the normalized repair bandwidth equals α and a precise determination of the file size of the new code is possible. FR codes from Steiner systems and their transposes, form a large class of base FR codes that satisfy these requirements. We

also demonstrate that the Kronecker product technique yields infinite families of FR codes that cannot be obtained from trivial β -expansion method. Furthermore, a careful analysis of the construction also allows to conclude that the failure resilience of these codes is as high as possible. We conclude by showing that the property of being resolvable is maintained under taking Kronecker products.

We begin with a simple example that generates a code that meets the Singleton bound. Let $\theta = 2a + 1$ for $a \geq 1$ and the incidence matrices N_1 and N_2 be equal to $J - I$ where J denotes $\theta \times \theta$ all-ones matrix and I denotes the identity matrix of the appropriate size. Then, the FR code \mathcal{C} obtained from the incidence matrix $\bar{N} = N_1 \otimes N_2$ has the following properties:

- The parameters of the code are $\bar{n} = \bar{\theta} = (2a + 1)^2$ and $\bar{\alpha} = \bar{\rho} = (2a)^2$.
- A failed node can be recovered by contacting two nodes.
- Contacting any two nodes recovers at least $2a(2a + 1)$ symbols. Thus, when $k = 2$, we have that the file size $\mathcal{M} = 2a(2a + 1)$, where it can be observed that $\lceil \frac{\mathcal{M}}{\bar{\alpha}} \rceil = 2$, so that the code meets the Singleton bound.

Example 13: Let $\mathcal{C} = (\Omega, V)$ be a FR code with $\Omega = \{1, 2, 3\}$ and $V = \{V_1 = \{2, 3\}, V_2 = \{1, 3\}, V_3 = \{1, 2\}\}$,

so that its incidence matrix $N = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. The new code

is obtained from the incidence matrix of $\bar{N} = N \otimes N$ and the storage nodes are shown in Fig. 7.

Suppose that the outer MDS code has parameters $(9, 6)$, so that $\theta = 9, \mathcal{M} = 6$. In this construction, the file can be recovered by contacting any two nodes, so that $k = 2$ and that a failed node can be recovered by contacting two nodes and downloading two packets from each of them.

Observation 3 (Non-trivial FR Codes with $\beta > 1$ obtained from Kronecker product): A FR code \mathcal{C} with parameters $(n, \theta, \alpha, \rho)$, yields a new FR code $\bar{\mathcal{C}}$ with parameters $(n^2, \theta^2, \alpha^2 \rho^2)$ via Kronecker product method with itself. If α does not divide θ then storage nodes of $\bar{\mathcal{C}}$ cannot be obtained from a trivial β -expansion with $\beta = \alpha$.

Example 14: Consider the FR code obtained by the Kronecker product of the Fano plane (shown in Fig. 2) with itself. The resultant code will have 49 symbols with nodes with storage capacity 9. If this code could be obtained by trivial

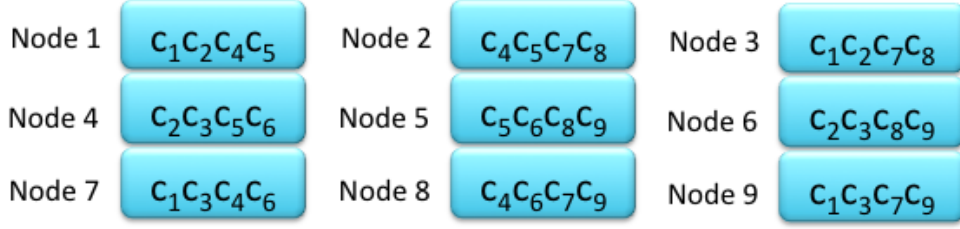


Fig. 7: A failed node can be recovered by contacting two nodes and downloading two packets from each of them. The code is resilient up to a total of three failures (corresponding to its minimum distance) and the file size is 6.

β -expansion from a base code with number of symbols $\tilde{\theta}$ and storage capacity $\tilde{\alpha}$, then there has to exist an integer m so that

$$\begin{aligned}\tilde{\theta}m &= 49, \text{ and} \\ \tilde{\alpha}m &= 9.\end{aligned}$$

As $9 \nmid 49$, the only feasible solution to the above system of equation is $\tilde{\theta} = 49$, $\tilde{\alpha} = 9$ and $m = 1$, which corresponds to the Kronecker product code.

In fact, there exists a family of codes whose parameters cannot be obtained via trivial β -expansion, as discussed in the corollary below.

Corollary 2: Let \mathcal{C} be a FR code obtained from a Steiner system $S(2, 3, 6u + 1)$ for some integer u . Then the FR code $\tilde{\mathcal{C}}$, which is obtained by the Kronecker product of \mathcal{C} with itself, cannot be obtained by trivial β -expansion with $\beta = 3$.

Lemma 10: Let $\mathcal{C}_1 = (\Omega_1, V_1)$ and $\mathcal{C}_2 = (\Omega_2, V_2)$ be two FR codes with parameters $(n_1, \theta_1, \alpha, \rho_1)$ and $(n_2, \theta_2, \alpha, \rho_2)$ such that any two storage nodes in \mathcal{C}_1 (or \mathcal{C}_2) have at most one symbol in common. Let \mathcal{M}_1 and \mathcal{M}_2 denote the file sizes of \mathcal{C}_1 and \mathcal{C}_2 respectively for a given $k_1 \leq \min\{n_1, n_2\}$. Suppose that either \mathcal{M}_1 or \mathcal{M}_2 is equal to $k_1\alpha - \binom{k_1}{2}$. Then the FR code \mathcal{C} obtained from Kronecker product of \mathcal{C}_1 and \mathcal{C}_2 has parameters $(n = n_1n_2, \theta = \theta_1\theta_2, \alpha^2, \rho_1\rho_2)$. The file size for \mathcal{C} when $k = k_1$ is given by $k_1\alpha^2 - \alpha\binom{k_1}{2}$.

Proof: Let N_1 and N_2 denote the incidence matrices of the FR codes \mathcal{C}_1 and \mathcal{C}_2 . Let c_i denote a column in N_1 and d_i denote a column in N_2 . The overlap between any two columns in $N_1 \otimes N_2$ can be expressed as $(c_i \otimes d_j)^t (c_{i'} \otimes d_{j'}) = c_i^t c_{i'} \otimes d_j^t d_{j'} \leq \alpha$. Thus the overlap between any two columns in $N_1 \otimes N_2$ is at most α and therefore the file size of \mathcal{C} is at least $k_1\alpha^2 - \alpha\binom{k_1}{2}$.

We know that any two nodes in \mathcal{C}_1 and \mathcal{C}_2 have at most one symbol in common. Thus, using a simple inclusion-exclusion principle argument implies that $\mathcal{M}_i \geq k_1\alpha - \binom{k_1}{2}$ for $i = 1, 2$. Furthermore, we are given that one of them meets this lower bound. Without loss of generality we assume that $\mathcal{M}_1 = k_1\alpha - \binom{k_1}{2}$. This implies that there exists a set of column vectors $\mathcal{I}_1 = \{c_1, \dots, c_{k_1}\}$ in N_1 such that they cover $\mathcal{M}_1 = k_1\alpha - \binom{k_1}{2}$ symbols, i.e., any two columns from \mathcal{I}_1 have exactly one symbol in common and any three columns from \mathcal{I}_1 have no symbols in common (see Appendix).

Next, we demonstrate a set of columns in $N_1 \otimes N_2$ that meets this lower bound. Let us consider a column in N_2 , denoted d_1 and examine $N_1 \otimes d_1$. Within this set we have

a subset of k_1 columns denoted $\mathcal{I}_2 = \{c_i \otimes d_1, \text{ for } c_i \in \mathcal{I}_1\}$. Now $(c_i \otimes d_1)^t (c_j \otimes d_1) = c_i^t c_j \otimes d_1^t d_1 = \alpha$, whereas any three column vectors from \mathcal{I}_2 will have a zero overlap. Thus, the number of symbols covered by this set is exactly $k_1\alpha^2 - \alpha\binom{k_1}{2}$.

This lemma can be used to determine the file size for the Kronecker product of certain Steiner systems.

Lemma 11: Let \mathcal{C} be a FR code obtained from a Steiner system $S(2, \alpha, \theta)$ with $\rho = \frac{\theta-1}{\alpha-1}$ such that it has a maximal arc of size $\rho+1$. Then the Kronecker product of the transposed code \mathcal{C}^T with itself is such that the file size equals $k\rho^2 - \rho\binom{k}{2}$ for $1 \leq k \leq \rho$.

Proof: The result follows from Lemma 3 and Lemma 10.

Remark 6: By Skolem's construction [35] we have $S(2, 3, \theta)$ for all $\theta \geq 7$ and $\theta \equiv 1, 3 \pmod{6}$. Moreover, for all $\theta \geq 7$ and $\theta \equiv 3, 7 \pmod{12}$ a Steiner system $S(2, 3, \theta)$ has at least one maximal arc [32]. Thus, Lemma 11 applies.

Lemma 12: Let N_1 and N_2 be incidence matrices of two FR codes such that the size of the pairwise intersection of distinct nodes is at most 1. Let $(n_1, \theta_1, \alpha, \rho_1)$ and $(n_2, \theta_2, \alpha, \rho_2)$ be parameters of these FR codes respectively. Assume that the FR code obtained from $\tilde{N} = N_1 \otimes N_2$ has normalized repair bandwidth $\beta = \alpha$. Then the FR code \tilde{N} is resilient up to $\rho_1\rho_2 - 1$ failures.

Proof: Define $\mathcal{N}(c_i)$ ($\mathcal{N}(d_j)$) to be the set of storage nodes in N_1 (N_2) that have exactly one symbol in common with c_i (d_j). As N_1 and N_2 are Steiner systems, two nodes have at most one symbol in common. In the discussion below we show that if there are at most $\rho_1\rho_2 - 1$ failures, we can recover all the nodes. We proceed by contradiction, i.e., assume that there exists a set of failed nodes F^* in \tilde{N} with $|F^*| = \rho_1\rho_2 - 1$. Suppose that there is a failed node $c_i \otimes d_j \in F^*$ that cannot be recovered. Note that $\beta = \alpha$. Thus, we need to download α symbols each from the surviving nodes, i.e., we need to consider nodes in \tilde{N} that have an overlap of α with $c_i \otimes d_j$.

Our first observation is that only the nodes in $\mathcal{N}(c_i) \otimes d_j$ and $c_i \otimes \mathcal{N}(d_j)$ are useful for recovering $c_i \otimes d_j$. To see this consider a node $c'_i \otimes d'_j$ in \tilde{N} such that it does not belong to $\mathcal{N}(c_i) \otimes d_j$ or $c_i \otimes \mathcal{N}(d_j)$. If $c'_i = c_i$, then $d'_j \notin \mathcal{N}(d_j)$, i.e., $(c'_i \otimes d'_j)^t (c_i \otimes d_j) = 0$; a similar argument holds when $c'_i \notin \mathcal{N}(c_i)$, $d'_j = d_j$. Otherwise $(c'_i \otimes d'_j)^t (c_i \otimes d_j)$ can be at most 1. Thus, only the nodes in $\mathcal{N}(c_i) \otimes d_j$ and $c_i \otimes \mathcal{N}(d_j)$ are useful for reconstructing $c_i \otimes d_j$.

Next, note that c_i (d_j) can be expressed as the sum of α

unit vectors of length θ_1 (θ_2). Let e_k denote the unit vector with a one in the k -th location. Thus, $c_i = \sum_{k \in I_1} e_k$, where $I_1 \subset [\theta_1]$ and $d_j = \sum_{l \in I_2} e_l$ where $I_2 \subset [\theta_2]$. Thus, the overlap between $c_i \otimes d_j$ and $\mathcal{N}(c_i) \otimes d_j$ can be expressed as $e_k \otimes d_j$ for some $k \in I_1$. A similar statement holds for the overlap between $c_i \otimes d_j$ and $c_i \otimes \mathcal{N}(d_j)$. Our next observation is that when we reconstruct $c_i \otimes d_j$, we can either download symbols from $\mathcal{N}(c_i) \otimes d_j$ or from $c_i \otimes \mathcal{N}(d_j)$ but not both. Indeed, for $k \in I_1, l \in I_2$, we have $(c_i \otimes e_l)^t (e_k \otimes d_j) = 1$. Thus, if we download symbols from both $\mathcal{N}(c_i) \otimes d_j$ and from $c_i \otimes \mathcal{N}(d_j)$, then we will need to download strictly more than α^2 symbols for reconstructing $c_i \otimes d_j$.

Note that there are ρ_1 copies of each $e_k \otimes d_j$, where $k \in I_1$. If there is at least one copy of $e_k \otimes d_j$, for all $k \in I_1$ available in the surviving nodes, then it is clear that $c_i \otimes d_j$ can be recovered by downloading copies of each $e_k \otimes d_j$ from the surviving nodes. Likewise, there are ρ_2 copies of each $c_i \otimes e_l$ for $l \in I_2$ and $c_i \otimes d_j$ can be recovered if each of these copies is available in the surviving nodes. In the discussion below we say that $c_i \otimes d_j$ is recoverable if either or both of these situations apply.

Thus, it is clear that if $c_i \otimes d_j$ is not recoverable it has to be the case that all copies of $e_{k^*} \otimes d_j$ for some $k^* \in I_1$ are unavailable. This implies that there exists a set of failed nodes denoted $F_1 \subset \mathcal{N}(c_i) \otimes d_j$ of size at least $\rho_1 - 1$. Arguing in a similar vein, we can consider whether $c_i \otimes d_j$ can be recovered from the nodes in $c_i \otimes \mathcal{N}(d_j)$. Based on the discussion above, if $c_i \otimes d_j$ is not recoverable, it has to be the case that there exists a set of failed nodes $F_2 \subset c_i \otimes \mathcal{N}(d_j)$ of size at least $\rho_2 - 1$. In addition the node sets $\mathcal{N}(c_i) \otimes d_j$ and $c_i \otimes \mathcal{N}(d_j)$ are disjoint, thus $F_1 \cap F_2 = \emptyset$, i.e., it is clear that at least $\rho_1 + \rho_2 - 2$ failures are essential to ensure that $c_i \otimes d_j$ is not recoverable.

Next, we examine whether any of the nodes in $F_1 \cup F_2$ are recoverable. A given node in F_1 is of the form $c_{i'} \otimes d_j$ where $c_{i'}^t c_{i'} = 1$. It is evident that $c_{i'} \otimes d_j$ cannot be recovered from $\mathcal{N}(c_{i'}) \otimes d_j$ as all copies of $e_{k^*} \otimes d_j$ for a specific k^* are unavailable owing to the failure of the nodes in F_1 . Specifically, note that it rules out the possibility of using the surviving nodes in the set $\mathcal{N}(c_i) \otimes d_j$. From the previous observation, it can only be recovered exclusively from the nodes in $c_{i'} \otimes \mathcal{N}(d_j)$.

Thus, there need to be at least $\rho_2 - 1$ failures from the node set $c_{i'} \otimes \mathcal{N}(d_j)$ to ensure that $c_{i'} \otimes d_j$ is not recoverable. Furthermore, these failures are distinct from the failures in $F_1 \cup F_2$. Arguing in this way for each node in F_1 , we conclude that at least $(\rho_1 - 1)(\rho_2 - 1)$ failures need to be induced to ensure that none of the nodes in F_1 can be recovered.

However, this implies a total of $1 + \rho_1 + \rho_2 - 2 + (\rho_1 - 1)(\rho_2 - 1) = \rho_1 \rho_2 > \rho_1 \rho_2 - 1$ failures. Thus, we conclude that even if an appropriate $F_1 \cup F_2$ can be found for $c_i \otimes d_j$, at least one node in F_1 can be recovered. After this recovery, the set F_1 cannot exist. This implies that $c_i \otimes d_j$ can be recovered. As the choice of $c_i \otimes d_j$ was arbitrary, we can recover any node when there are at most $\rho_1 \rho_2 - 1$ failures.

This bound is tight since each symbol in \bar{N} is repeated $\rho_1 \rho_2$ times. Thus, we can easily find a set of $\rho_1 \rho_2$ failures that we cannot recover from.

Corollary 3: Let N_1 and N_2 be transposes of incidence matrices of two Steiner systems namely $S(2, \alpha_1, \theta_1)$ and $S(2, \alpha_2, \theta_2)$ where the parameters satisfy $\rho = \frac{\theta_1 - 1}{\alpha_1 - 1} = \frac{\theta_2 - 1}{\alpha_2 - 1}$. Assume the FR code obtained from \bar{N} has normalized repair bandwidth $\beta = \rho$. Then, the FR code is resilient up to $\alpha_1 \alpha_2 - 1$ failures.

Proof: Any two nodes meet in exactly one symbol in the FR code obtained by transposes of incidence matrices of a Steiner system. Also note that the main ingredient of the proof Lemma 12 is the property that two nodes meet in at most one symbol in Steiner systems. So the rest follows similarly as in the previous proof.

We also investigate the properties of FR codes that are generated by taking the Kronecker product of net FR codes with themselves. The Kronecker product does not necessarily produce a new net FR code but it yields a resolvable FR code. For example, in Fig. 8 a resolvable FR code is obtained from the Kronecker product of a net FR code with itself. However, the obtained code is not a net FR code. To see this, we note that that node sets $\{1, 5, 9, 13\}$ and $\{2, 6, 10, 14\}$ form parallel classes, but the intersection sizes of node 1 with the nodes in the set $\{2, 6, 10, 14\}$ are either two or zero, which implies that the obtained code is not a net FR code.

Lemma 13: Let N be the incidence matrix of a net FR code with parameters $(n, \theta, \alpha, \rho)$. Then, the FR code obtained from $\bar{N} = N \otimes N$ is a resolvable FR code.

Proof: We can order the columns of N with respect to the ρ parallel classes. Assume that the j -th block in i -th parallel class is represented by the column $c_{i,j}$. We will show for fixed i and s , $c_{i,j} \otimes c_{s,r}$ with $1 \leq j \leq \frac{\theta}{\alpha}$ and $1 \leq r \leq \frac{\theta}{\alpha}$ forms a set of blocks which is a parallel class. There will be $\frac{\theta^2}{\alpha^2}$ blocks in this set, hence it is enough to show any distinct two blocks does not share any points. Since $(c_{i,j}^t \otimes c_{s,r}^t)(c_{i,u} \otimes c_{s,v}) = (c_{i,j}^t c_{i,u} \otimes c_{s,r}^t c_{s,v})$ equals the zero, the $\frac{\theta^2}{\alpha^2}$ vectors form a parallel class.

Example 15: A simple example can be obtained from $\mathcal{C} = (\Omega, V)$ where $\Omega = \{1, 2, 3, 4\}$ and $V = \{V_1 = \{1, 2\}, V_2 = \{3, 4\}, V_3 = \{1, 3\}, V_4 = \{2, 4\}\}$. The code obtained from $\bar{N} = N \otimes N$ is illustrated in Fig. 8.

V. CONSTRUCTION OF FR CODES WHEN $d < k$

In the discussion so far, we have considered FR codes where the recovery degree $d \geq k$, i.e., the repair degree (d) of the code is at least as high as the number of nodes (k) contacted for recovering the file. Of course, the codes operate at the MBR point which implies that they download exactly α symbols for regeneration. However, as discussed in Section I, in many application scenarios it has been recognized that the number of nodes that the new node has to contact is an important metric that needs to be optimized, rather than the repair bandwidth. Note that the definition of a FR code does not rule out codes where $d < k$.

In this section, we discuss constructions of *locally recoverable FR codes* that have the property that $d < k$. It turns out that the minimum distance bound for locally recoverable codes that was derived in [3], [4], needs to be refined for our

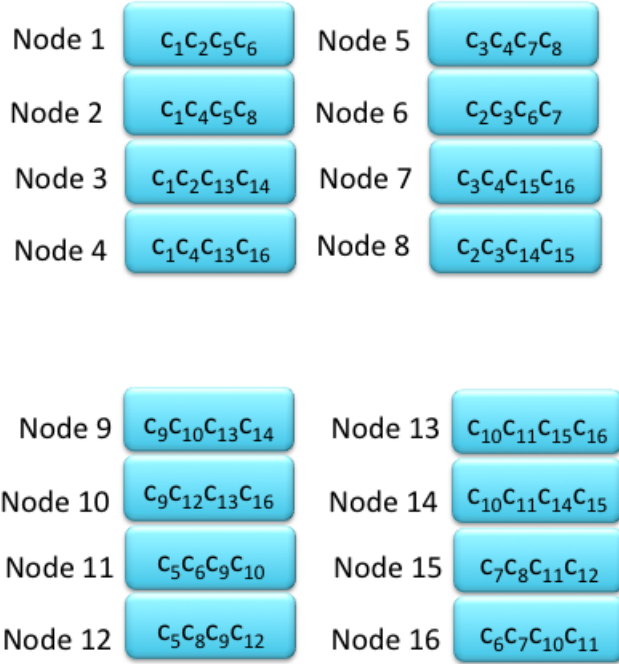


Fig. 8: The resultant FR code has $\Omega = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9, c_{10}, c_{11}, c_{12}, c_{13}, c_{14}, c_{15}, c_{16}\}$. Each storage node contains 4 symbols. A failed node can be recovered by contacting two nodes and downloading 2 packets from each. The code is resilient up to 3 failures.

scenario of exact, uncoded and table-based repair. We derive such a bound and present constructions that meet this bound.

Definition 17 (Locally recoverable fractional repetition code): Let $\mathcal{C} = (\Omega, V)$ be a FR code for a (n, k, d, α) -DSS, with repetition degree ρ and normalized repair bandwidth $\beta = \alpha/d$. If the repair degree $d < k$, then the FR code \mathcal{C} is called a locally recoverable fractional repetition code.

As before we define ρ_{res} to be the maximum number of node failures such that each failed node can be recovered by contacting d surviving nodes and downloading symbols from them. For a node $V_i \in V$ in \mathcal{C} , let $\mathcal{S}(V_i) \in V$ denote the set of nodes (with $|\mathcal{S}(V_i)| < k$) that are contacted if V_i fails. We refer to $\mathcal{S}(V_i)$ as the local structure associated with V_i . Note that it is possible that the set of nodes in $\mathcal{S}(V_i)$ and the corresponding symbols form a FR code (cf. Definition 2); however this is not essential.

A. Codes for systems with $\rho_{res} = 1$

Our first construction is a class of codes which is optimal with respect to the bound provided in Lemma 2 and allow local recovery in the presence of a single failure. Our construction leverages the properties of undirected graphs with large girth². The basic idea is to associate the edges of the undirected graph with the symbols and the vertices with the storage nodes. Each storage node stores its incident symbols. We explain this

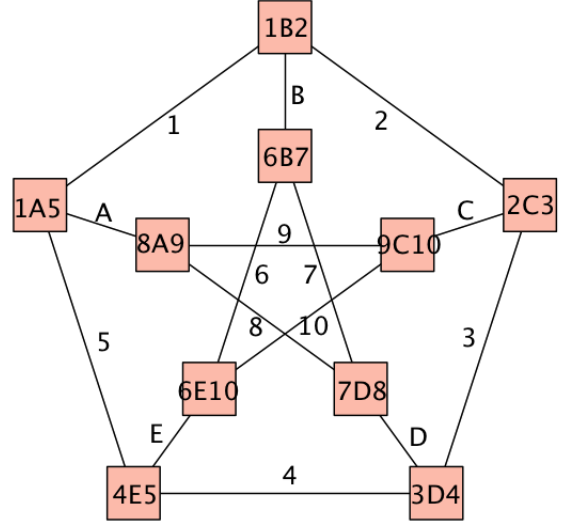


Fig. 9: The figure shows the Petersen graph with its edges labeled from 1, ..., 10 and A, ..., E. Each vertex acts as a storage node and stores the symbols incident on it.

construction and highlight the intuition behind it by means of the following example.

Example 16: The Petersen graph on 10 vertices and 15 edges is a 3-regular graph with girth 5. We label the edges 1, ..., 10 and A, B, ..., E in Fig. 9. If a given storage node fails, it is evident that it can be regenerated by contacting its corresponding neighbors in the Petersen graph and downloading one symbol each from them. For instance, if node $\{1, A, 5\}$ fails, it can download one symbol each from $\{1, B, 2\}$, $\{8, A, 9\}$ and $\{4, E, 5\}$. Next, note that there is no cycle of length 4, in the Petersen graph. Thus, if we consider any collection of four nodes (as an example), we are guaranteed that the number of edges incident on them is reasonably large. This allows to assert that the file size for such $k = 4$ is high. In fact, in the subsequent discussion we show that the file size in this case and for $k = 5$ meets the minimum distance bound for locally recoverable codes.

We now formalize the basic intuition in the above example, by considering general graphs and precisely calculating the file sizes and minimum distance bounds.

Definition 18: An undirected graph Γ is called an (s, g) -graph if each vertex has degree s , and the length of the shortest cycle in Γ is g .

Construction 1: Let $\Gamma = (V', E')$ be a (s, g) -graph with $|V'| = n$.

- (i) Arbitrarily index the edges of Γ from 1 to $\frac{ns}{2}$.
- (ii) Each vertex of Γ corresponds to a storage node and stores the symbols incident on it.

The above procedure yields a FR code $\mathcal{C} = (\Omega, V)$ with n storage nodes, parameters $\theta = \frac{ns}{2}$, $\alpha = s$ and $\rho = 2$. Upon single failure, the failed node can be regenerated by downloading one symbol each from the storage nodes corresponding to the vertices adjacent to it in Γ (i.e., $\beta = 1$); thus, the repair degree $d = s$. Note that for this construction, the local

²The girth of a graph is the length of its shortest cycle.

structures are typically not FR codes. Suppose that the storage node corresponding to vertex $v_1 \in \Gamma$ fails, then we contact the storage nodes corresponding to its $(s-1)$ neighbors in Γ ; this is the local structure associated with v_1 . If the girth $g > 3$, then it is clear that the nodes in the local structure do not have symbols in common, i.e., they do not form a FR code.

We note that the work of [7] also used the above construction for MBR codes; however, they did not have the girth restriction on Γ . As we discuss next, (s, g) -graphs allow us to construct locally recoverable codes and provide a better bound on the file size when $k \leq g$. We allow the system parameter k to be greater than d , however in the work of [7], they consider only the case $k \leq d$. The work of [30] also used high-girth graphs, but their constructions are not in the context of locally recoverable codes.

Lemma 14: Let $\mathcal{C} = (\Omega, V)$ be a FR code constructed by Construction 1. If $s > 2$, and $k \leq g$, we have $|\cup_{i=1}^k V_i| \geq k(s-1)$ for any $V_i \in V, i = 1, \dots, k$.

Proof: Let V_1, V_2, \dots, V_{k-1} and V_k be any k nodes in our DSS, where $k \leq g$. We argue inductively. Note that $|V_1| = s > s-1$. Suppose that $|\cup_{i=1}^j V_i| \geq j(s-1) + \xi$ for $j < k$, where $1 \leq \xi \leq j$ is the number of connected components formed by the nodes V_1, \dots, V_j in Γ . Now consider $|\cup_{i=1}^{j+1} V_i|$ where $j+1 < k$. Note that since $j+1 < g$ there can be no cycle in $\cup_{i=1}^{j+1} V_i$. Thus, V_{j+1} is connected at most once to each connected component in $\cup_{i=1}^j V_i$. Suppose that V_{j+1} is connected to ℓ existing connected components in $\cup_{i=1}^j V_i$, where $0 \leq \ell \leq \min(\xi, s)$. Then, the number of connected components in $\cup_{i=1}^{j+1} V_i$ is $\xi - \ell + 1$ and the number of new symbols that it introduces is $s - \ell$. Therefore $|\cup_{i=1}^{j+1} V_i| = j(s-1) + \xi + s - \ell = (j+1)(s-1) + \xi - \ell + 1$. This proves the induction step.

Thus, $|\cup_{i=1}^{k-1} V_i| \geq (k-1)(s-1) + \xi_{k-1}$, where ξ_{k-1} is the number of connected components formed by V_1, \dots, V_{k-1} . Now consider $\cup_{i=1}^k V_i$. Note that there can be a cycle introduced at this step if $k = g$. Now, if $\xi_{k-1} \geq 2$, it can be seen that V_k can only connect to each of the ξ_{k-1} connected components once, otherwise it would imply the existence of a cycle of length strictly less than g in Γ . Thus, in this case $|\cup_{i=1}^k V_i| \geq k(s-1)$. On the other hand if $\xi_{k-1} = 1$, then V_k can connect at most twice to this connected component. In this case again we can observe that $|\cup_{i=1}^k V_i| \geq k(s-1)$.

Lemma 15: Let $\Gamma = (V, E)$ be a (s, g) -graph with $|V| = n$ and $s > 2$. If $g \geq k = as + b$ such that $s > b \geq a + 1$, then \mathcal{C} obtained from Γ by Construction 1 is optimal with respect to the minimum distance bound in Lemma 2 when the file size $\mathcal{M} = k(s-1)$.

Proof: We have

$$k(s-1) = (as+b)(s-1) = as^2 + (b-a)s - b.$$

Since, $s > b \geq a + 1$ the following holds.

$$\left\lceil \frac{k(s-1)}{s} \right\rceil = \left\lceil \frac{as^2 + (b-a)s - b}{s} \right\rceil = as + (b-a),$$

and

$$\left\lceil \frac{k(s-1)}{s^2} \right\rceil = \left\lceil \frac{as^2 + (b-a)s - b}{s^2} \right\rceil = \left\lceil a + \frac{(b-a)s - b}{s^2} \right\rceil = \left\lceil a + \frac{b-a}{s} \right\rceil = a + 1$$

From Lemma 14, any k nodes cover at least $k(s-1)$ symbols. Thus, the code is minimum distance optimal since

$$\left\lceil \frac{k(s-1)}{s} \right\rceil + \left\lceil \frac{k(s-1)}{s^2} \right\rceil = k + 1. \text{ (cf. Observation 1)}$$

Corollary 4: Let $\Gamma = (V, E)$ be a (s, g) -graph with $|V| = n$ and $s > 2$. If $g \geq s + 2$, then \mathcal{C} obtained from Γ by Construction 1 is optimal with respect to the bound in Lemma 2 for file size $\mathcal{M} = s^2 + s - 2$.

It can be observed that in the specific case of $s = 2$, applying Construction 1 results in a DSS where the union of any k nodes has at least $k + 1$ symbols. We now discuss some examples of codes that can be obtained from our constructions.

Sachs [39] provided a construction which shows that for all $s, g \geq 3$, there exists a s -regular graph of girth g . Also, explicit constructions of graphs with arbitrarily large girth are known [40]. Using these we can construct infinite families of optimal locally recoverable codes.

An (s, g) -graph with the fewest possible number of vertices, among all (s, g) -graphs is called an (s, g) -cage and will result in the maximum code rate for our construction. For instance, the $(3, 5)$ -cage is the Petersen graph. We note here that bipartite cages of girth 6 were used to construct FR codes in [28] though these were not in the context of locally recoverable codes. An exhaustive survey of cages can be found in [13].

B. Codes for systems with $\rho_{res} > 1$

Our second class of codes are such that the local structures are also FR codes. The primary motivation for considering this class of codes is that they naturally allow for local recovery in the presence of more than one failure as long as the local FR code has a repetition degree greater than two. Thus, in these codes, each storage node participates in one or more local FR codes that allow local recovery in the presence of failures. We motivate the design of these FR codes by means of the following example.

Example 17: An example of such a code is shown in Fig. 10. The main idea is to have four FR codes derived from the Fano plane that are supported on disjoint sets of symbols. We refer to each of these FR codes as local structures. Note that if there are at most two failures, the nodes can be regenerated by simply downloading symbols from the corresponding local structures. Moreover, upon inspection, it is not too hard to see that any set of 15 nodes cover at least 17 symbols. Thus, we obtain an instance of a local FR code with $n = \theta = 28, \alpha = 3, \rho = 3$ that has $k = 15$ and $d = 3$. As $d < k$, this FR code is local.

Note that it is relatively easy to obtain local codes in such a manner, i.e., by considering a collection of FR codes supported on disjoint sets of symbols. However, one really needs to measure them with respect to minimum distance bound for local codes. We did this evaluation for the codes from high girth graphs presented above (cf. Lemma 15) and demonstrated that for certain ranges of k , the constructed codes were minimum distance optimal. However, we emphasize the minimum distance bound for local codes in Lemma 2 holds for general codes. In our class of codes, we have the added



Fig. 10: The figure shows a DSS where $n = 28, k = 15, r = 3, \theta = 28, \alpha = 3, \rho = 3$ and each local FR code (the columns in the figure) is a projective plane of order 2 which is also known as a Fano plane. Here, $\rho^{res} = 2$. Any set of 15 nodes cover at least $\mathcal{M} = 17$ symbols. Thus, the minimum distance of the code is 14 when the file size $\mathcal{M} = 17$.

requirement that each node participates in a local structure that allows it to be recovered by download in case of failure. Accordingly the bound in Lemma 2 is too loose.

For the class of codes that we consider, we derive an upper bound on the minimum distance of such codes when the file size is larger than the number of symbols in one local structure. Following this, we examine (fairly technical) conditions on the local structures that in turn allow for minimum distance optimality of the local FR code. We also demonstrate that several FR codes satisfy these conditions and conclude with some example of minimum distance optimal local FR codes.

Lemma 16: Let \mathcal{C} be a locally recoverable FR code with parameters $(n, \theta, \alpha, \rho)$ where each node belongs to a local FR code with parameters $(n_{loc}, \theta_{loc}, \alpha, \rho_{loc})$. Suppose that the file size $\mathcal{M} > \theta_{loc}$. Then,

$$d_{min} \leq \max \left(n - \left\lceil \frac{\mathcal{M}\rho_{loc}}{\alpha} \right\rceil + \rho_{loc}, n + n_{loc} + 1 - \left\lceil \frac{\mathcal{M}\rho_{loc} + \theta_{loc}}{\alpha} \right\rceil \right).$$

Proof:

We will apply an algorithmic approach here (inspired by the one used in [3]). Namely, we iteratively construct a large enough set $\mathcal{S} \subset V$ so that $|\mathcal{S}| < \mathcal{M}$. The minimum distance bound is then given by $n - |\mathcal{S}|$. Our algorithm is presented in Fig. 11. Towards this end, let S_i and $H(S_i)$ represent the number of nodes and the number of symbols included at the end of the i -th iteration. Furthermore, let $s_i = |S_i| - |S_{i-1}|$ and $h_i = |H(S_i)| - |H(S_{i-1})|$, represent the corresponding increments between the $(i-1)$ -th and the i -th iteration. We divide the analysis into two cases.

- **Case 1:** [The algorithm exits without ever entering line 8.] Note that we have $1 \leq s_i \leq n_{loc}$ and $h_i \leq \theta_{loc} -$

```

1:  $S_0 = \emptyset, i = 1$ 
2: while  $H(S_{i-1}) < \mathcal{M}$  do
3:   For each node  $Y_j \in S_{i-1}$ , identify a FR code  $Pf_j = (\Omega_{Pf_j}, V_{Pf_j})$  (if it exists) such that  $Y_j \in V_{Pf_j}, V_{Pf_j} \not\subseteq S_{i-1}$ . If no such FR code exists, find a FR code that has no intersection with  $S_{i-1}$  and set  $Pf_1$  equal to it.
   • Let  $b_j = |\Omega_{Pf_j} \cap H(S_{i-1})|$ . Let  $j^* = \arg \max_j b_j$ .
4:   if  $\theta_{loc} - b_{j^*} + H(S_{i-1}) < \mathcal{M}$  then
5:     Set  $S_i = S_{i-1} \cup V_{Pf_{j^*}}$ .
6:   else
7:     if there exists  $A \subset V_{Pf_{j^*}}$  such that  $|S_{i-1} \cup A| > |S_{i-1}|$  and  $H(S_{i-1} \cup A) < \mathcal{M}$  then
8:       Let  $V_{Pf'_{j^*}} = \arg \max_{A \subset V_{Pf_{j^*}}} H(S_{i-1} \cup A) < \mathcal{M}$ . Set  $S_i = S_{i-1} \cup V_{Pf'_{j^*}}$ .
9:   else
10:    Exit.
11:   end if
12: end if
13: end while

```

Fig. 11: Algorithm for finding the distance bound

$a(n_{loc} - s_i)$ where $a(n_{loc} - s_i)$ is the minimum number of symbols covered by $(n_{loc} - s_i)$ nodes in the local FR code and hence a lower bound on $|\Omega_{Pf_{j^*}} \cap H(S_{i-1})|$. By considering the bipartite graph representing the local FR code (cf. Definition 8) We see that $a(n_{loc} - s_i) \geq \frac{(n_{loc} - s_i)\alpha}{\rho_{loc}}$. Thus, we have

$$\theta_{loc} - a(n_{loc} - s_i) \leq \theta_{loc} - \frac{n_{loc}\alpha - s_i\alpha}{\rho_{loc}} = \frac{s_i\alpha}{\rho_{loc}}.$$

Suppose that the algorithm runs for l iterations and exits on the $l + 1$ iteration. Then

$$\sum_{i=1}^l s_i \geq \frac{\rho_{loc}}{\alpha} \sum_{i=1}^l h_i.$$

Since the algorithm exits without ever entering line 8, it is unable to accumulate even one additional node. Hence

$$\begin{aligned} \sum_{i=1}^l h_i &\geq \mathcal{M} - \alpha, \text{ which implies that} \\ \sum_{i=1}^l s_i &\geq \left\lceil \frac{\rho_{loc}}{\alpha} (\mathcal{M} - \alpha) \right\rceil \text{ by the integer constraint.} \end{aligned}$$

Thus, the bound on the minimum distance becomes

$$d_{min} \leq n - \left\lceil \frac{\rho_{loc} \mathcal{M}}{\alpha} \right\rceil + \rho_{loc}.$$

- **Case 2:** [The algorithm exits after entering line 8.] Note that by assumption, $\mathcal{M} > \theta_{loc}$. Suppose that the algorithm enters line 5, $l \geq 1$ times. Now we have $\sum_{i=1}^l h_i \geq \mathcal{M} - \theta_{loc}$, otherwise we could include another local structure. Hence we need to add nodes so that strictly less than $\mathcal{M} - \sum_{i=1}^l h_i$ symbols are covered. It can be observed

that we can include at least $\left\lceil \frac{\mathcal{M} - \sum_{i=1}^l h_i}{\alpha} \right\rceil - 1$ more nodes. Therefore, the total number of nodes accumulated is

$$\begin{aligned} &\geq \frac{\rho_{loc}}{\alpha} \sum_{i=1}^l h_i + \left\lceil \frac{\mathcal{M} - \sum_{i=1}^l h_i}{\alpha} \right\rceil - 1 \\ &\geq \frac{\rho_{loc} - 1}{\alpha} (\mathcal{M} - \theta_{loc}) + \frac{\mathcal{M}}{\alpha} - 1 \\ &= \frac{\mathcal{M} \rho_{loc} + \theta_{loc}}{\alpha} - n_{loc} - 1. \end{aligned}$$

Therefore, we have the following minimum distance bound.

$$d_{min} \leq n + n_{loc} + 1 - \left\lceil \frac{\mathcal{M} \rho_{loc} + \theta_{loc}}{\alpha} \right\rceil.$$

The final bound is obtained by taking the maximum of the two bounds obtained above.

The following corollary can be also be established.

Corollary 5: Let \mathcal{C} be a locally recoverable FR code with parameters $(n, \theta, \alpha, \rho)$ where each node belongs to a local FR code with parameters $(n_{loc}, \theta_{loc}, \alpha, \rho_{loc})$. Furthermore, suppose that \mathcal{C} can be partitioned as the union of ℓ disjoint local FR codes. If the file size $\mathcal{M} = t\theta_{loc} + \beta$ for some integer $1 \leq t < \ell$ and $\beta \leq \alpha$, we have $d_{min} \leq n - \left\lceil \frac{\mathcal{M} \rho_{loc}}{\alpha} \right\rceil + \rho_{loc}$.

Proof: Applying the algorithm in Fig. 11 it can be observed that we will never enter line 8, as \mathcal{C} consists of the union of disjoint local FR codes and the file size $\mathcal{M} = t\theta_{loc} + \beta$. Thus, after accumulating t disjoint local FR codes, the algorithm will exit, yielding the required bound.

Construction 2: Let $\mathcal{C} = (\Omega, V)$ be a FR code with parameters $(n, \theta, \alpha, \rho)$ such that any $\Delta + 1$ nodes in V cover θ symbols and for $V_i, V_j \in V$, we have $|V_i \cap V_j| \leq \beta$ when $i \neq j$. We construct a locally recoverable FR code $\bar{\mathcal{C}}$ by considering the disjoint union of $l (> 1)$ copies of \mathcal{C} . Thus, $\bar{\mathcal{C}}$ has parameters $(ln, l\theta, \alpha, \beta)$. We call \mathcal{C} the local FR code of $\bar{\mathcal{C}}$.

Lemma 17: Let $\bar{\mathcal{C}}$ be a code constructed by Construction 2 for some $l > 1$ such that the parameters of the local FR code satisfy $(\rho - 1)\alpha\theta - (\theta + \alpha)(\Delta - 1)\beta \geq 0$. Let the file size be $\mathcal{M} = t\theta + \alpha$ for some $1 \leq t < l$. Then $\bar{\mathcal{C}}$ is optimal with respect to Corollary 5.

Proof: It is evident that $\bar{\mathcal{C}}$ is the disjoint union of l local FR codes. Thus, the minimum distance bound here is $d_{min} \leq ln - \left\lceil \frac{(t\theta + \alpha)\rho}{\alpha} \right\rceil + \rho = (l - t)n$. The code is optimal when any $tn + 1$ nodes in $\bar{\mathcal{C}}$ cover at least $\mathcal{M} = t\theta + \alpha$ symbols. We show that this is the case below.

Let a_i be the number of nodes that are chosen from the i -th local FR code and X_i be the symbols covered by these a_i nodes. Note that for any $1 \leq i \leq l$ if $a_i \geq \Delta + 1$, then $X_i = \theta$ (the maximum possible). Suppose there are $0 \leq t_1 \leq t$ local FR codes that cover θ symbols. In this case it suffices to show that $(t - t_1)n + 1$ nodes cover at least $(t - t_1)\theta + \alpha$ symbols. Here we can omit case of $t = t_1$, since our claim clearly holds in this situation. Suppose that these nodes belong to s local FR codes, where $a_i \leq \Delta, i = 1, \dots, s$. By applying Corradi's lemma [41] we obtain

$$|X_i| \geq \frac{\alpha^2 a_i}{\alpha + (a_i - 1)\beta} \geq \frac{\alpha^2 a_i}{\alpha + (\Delta - 1)\beta}.$$

This implies that

$$\begin{aligned} \sum_{i=1}^s |X_i| &\geq \sum_{i=1}^s \frac{\alpha^2 a_i}{\alpha + (\Delta - 1)\beta} \\ &= \frac{\alpha^2}{\alpha + (\Delta - 1)\beta} \sum_{i=1}^s a_i \\ &= \frac{\alpha^2}{\alpha + (\Delta - 1)\beta} ((t - t_1)n + 1) \\ &= \frac{(t - t_1)\theta\rho\alpha}{\alpha + (\Delta - 1)\beta} + \frac{\alpha^2}{\alpha + (\Delta - 1)\beta} \text{ (since } n\alpha = \theta\rho) \\ &= (t - t_1)\theta + \left(\frac{\rho\alpha}{\alpha + (\Delta - 1)\beta} - 1 \right) (t - t_1)\theta + \frac{\alpha^2}{\alpha + (\Delta - 1)\beta} \\ &\geq (t - t_1)\theta + \frac{((\rho - 1)\alpha - (\Delta - 1)\beta)\theta + \alpha^2}{\alpha + (\Delta - 1)\beta} \\ &\geq (t - t_1)\theta + \alpha \text{ (using the assumed conditions).} \end{aligned}$$

The above lemma can be used to generate several examples of locally recoverable codes with $\rho_{res} > 1$. We discuss two examples below.

Example 18: Let q be a prime power. We consider the codes obtained from affine resolvable designs discussed in Section III-B1. These codes have parameters $\theta = q^m, \alpha = q^{m-1}, \rho = \frac{q^m - 1}{q - 1}$ and $n = q\rho$. These codes are resolvable and hence we can vary the repetition degree by choosing an appropriate number of parallel classes. Note that the number of nodes in a parallel class is $\theta/\alpha = q$.

Suppose we choose the local FR code by including q^{m-1} parallel classes, so that the repetition degree is q^{m-1} and there

are $n = q^m$ nodes. Furthermore, since the design is affine resolvable, $\beta = q^{m-2}$. The value of Δ (cf. Definition 2) can be determined as follows. For the local FR code, any subset of at least $q^m - q^{m-1} + 1$ nodes has at least one intact parallel class, which covers all the $\theta = q^m$ symbols. Accordingly, for this code we can conclude that $\Delta = q^m - q^{m-1}$.

Next, we verify the conditions of Lemma 17. For this local FR code, we have that

$$\begin{aligned} & (\rho - 1)\alpha\theta - (\theta + \alpha)(\Delta - 1)\beta \\ &= (q^{m-1} - 1)q^{2m-1} - (q^m + q^{m-1})(q^m - q^{m-1} - 1)q^{m-2} \\ &= q^{2m-3}(q^{m+1} - q^2 - (q + 1)(q^m - q^{m-1} - 1)) \\ &= q^{2m-3}(q^{m+1} - q^2 - (q^m - q^{m-1} - 1) - (q^{m+1} - q^m - q)) \\ &= q^{2m-3}(q^{m-1} + q + 1 - q^2) \\ &\geq 0, \text{ when } m \geq 3. \end{aligned}$$

Thus, to summarize for the local FR code under consideration, the conditions of Lemma 17 apply when $m \geq 3$. Thus, we can construct a FR code by consider the disjoint union of l of these local FR codes using Construction 2. The code will be optimal with respect to the bound derived in Corollary 5 for file sizes of the form $tq^m + q^{m-1}$ for $1 \leq t < l$.

Example 19: A projective plane of order q also forms a FR code $\mathcal{C} = (\Omega, V)$, where $\alpha = q + 1$ and $\rho = q + 1$. Furthermore, $|V_i \cap V_j| = 1$ if $i \neq j$ and each pair of symbols appears in exactly one node; this further implies that $\beta = 1$. A simple counting argument shows that $|\Omega| = \theta = q^2 + q + 1$ and $n = q^2 + q + 1$. The value of Δ (cf. Definition 2) can be determined in the following manner. Applying Corradi's Lemma, we note that any $q^2 + 1$ nodes cover at least a number of symbols greater than or equal to

$$\begin{aligned} \frac{(q+1)^2(q^2+1)}{q^2+q+1} &= q^2 + q + \frac{q+1}{q^2+q+1} \\ &> q^2 + q, \end{aligned}$$

whereby we conclude that $q^2 + 1$ nodes cover all the $q^2 + q + 1$ symbols. It can also be observed that there is a set of q^2 nodes that do not cover all the $q^2 + q + 1$ symbols as the repetition degree of the symbols is $q + 1$. Thus, in this case we can observe that $\Delta = q^2$.

We construct a locally recoverable FR code $\bar{\mathcal{C}}$ by taking $l > 1$ copies of the code \mathcal{C} . So the code $\bar{\mathcal{C}}$ has parameters $(l(q^2 + q + 1), l(q^2 + q + 1), q + 1, q + 1)$. Let the file size be $\mathcal{M} = t(q^2 + q + 1) + q + 1$ for some $1 \leq t < l$. Then, $\bar{\mathcal{C}}$ is optimal with respect to Lemma 16 and has $\rho_{res} = q$. An example is illustrated in Fig. 10.

VI. CONCLUSIONS AND FUTURE WORK

In this work we have constructed several classes of fractional repetition codes that can be used in distributed storage systems. These codes allow for a repair process that is exact and uncoded but table-based. Our constructions stem from combinatorial designs such as Steiner systems, affine geometries, Hadamard designs and mutually orthogonal Latin squares. We demonstrate that (i) the repetition degree of the symbols which dictates the failure resilience of the code can be varied in an easy manner, and (ii) construct instances of

codes with $\beta > 1$ that cannot be obtained in a trivial manner from codes with $\beta = 1$. In addition, we show that new FR codes can be obtained from taking Kronecker products of existing ones and analyze their properties. For codes with exact, uncoded and local repair property (where $d < k$), we establish an appropriate minimum distance bound and present constructions from high-girth graphs and collections of local FR codes (with specific properties) that meet these bounds. For most of our constructions, we determine the code rate for specific ranges of k .

There are several opportunities for future work. It would be interesting to examine applications of designs in other areas of network coding. For instance, [42] shows that designs can be used to construct directed acyclic networks that have nontrivial implications for distributed function computation. In principle, several combinatorial designs can be treated as FR codes. However, it would be interesting to examine if there are other families that have desirable properties and lend themselves to an analysis of the system code rate. It is to be noted that the code rate depends on the minimum size of the union of k -sized subsets of the storage nodes. It can also be viewed as determining the expansion level of a bipartite graph derived from the incidence matrix of the design. In general, it is somewhat challenging as most results in the literature only discuss pairwise intersections. A related problem would be determine feasible and infeasible parameter ranges for FR codes.

VII. ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers whose comments and suggestions significantly improved the quality of the paper.

VIII. APPENDIX

Proof of Lemma 4.

Note that the properties of net FR codes imply that any two storage nodes intersect in either one or zero symbols. Thus, $\alpha k - \binom{k}{2}$ is the lower bound on the file size. In the discussion below we demonstrate the existence of k nodes that cover exactly $\alpha k - \binom{k}{2}$ symbols. Let the parallel classes be indexed from 0 to $\rho - 1$.

- 1: Choose a node V_0 from 0-th parallel class. Initialize $H = \emptyset$, $S = \{V_0\}$ and $i = 1$.
- 2: **while** $|H| \leq a$ and $|S| < k$ **do**
- 3: Choose V_i from the i -th parallel class such that $V_\ell \cap V_i \notin H$ for all $V_\ell \in S$.
- 4: Set $H = H \cup \bigcup_{V_\ell \in S} V_\ell \cap V_i$.
- 5: Set $S = S \cup V_i$.
- 6: **end while**

We need to show that an appropriate V_i can always be chosen in the algorithm and that $|S| = k$ upon exit. To see this note that H tracks the set of pairwise intersections between the nodes at all times. At the beginning of stage i , the size of H is at most $\binom{i}{2}$ (by interpreting $\binom{1}{2} = 0$). Note that a parallel class has a nodes and that two nodes from the same parallel class do not intersect. Thus, as long as $a > \binom{i}{2}$ we can always

find an appropriate V_i . By our assumption $\binom{k-1}{2} < a$. Thus, the algorithm exits with $|S| = k$.

Lemma 18: Consider sets A_1, \dots, A_k such that $|A_i| = \alpha$ and $|A_i \cap A_j| \leq 1$ when $i \neq j$ and $|\cup_{i=1}^k A_i| = k\alpha - \binom{k}{2}$. This implies that $|A_i \cap A_j| = 1$ for $i \neq j$ and $|A_i \cap A_j \cap A_l| = 0$ for all distinct triples (i, j, l) where $i, j, l = 1, \dots, k$.

Proof: By the inclusion-exclusion principle, we have that

$$|\cup_{i=1}^k A_i| \geq \sum_i |A_i| - \sum_{i < j} |A_i \cap A_j| \geq k\alpha - \binom{k}{2}.$$

However, as $|\cup_{i=1}^k A_i| = k\alpha - \binom{k}{2}$, this implies that $|A_i \cap A_j| = 1$ for all pairs (i, j) such that $i \neq j$.

For a set $I \subseteq [k]$, let A_I denote the set $\cap_{i \in I} A_i$. We note that the given conditions also imply that

$$\sum_{\emptyset \neq I \subseteq [k], |I| \geq 3} (-1)^{|I|+1} |A_I| = 0. \quad (9)$$

We argue that it has to be the case that $|A_I| = 0$ for $\emptyset \neq I \subseteq [k]$, $|I| \geq 3$. Suppose that this is not the case and there are l subsets I_1, \dots, I_l such that $|I_i| \geq 3$, $i = 1, \dots, l$ and $|A_{I_i}| = 1$. For each I_i , there has to be a maximal I_i^* such that $I_i \subset I_i^*$. Moreover, it has to hold that $|I_i^* \cap I_j^*| \leq 1$, as otherwise $I_i^* \cup I_j^*$ provides an example of a subset that is larger than both I_i^* and I_j^* . This establishes that for each I_i , there is a *unique* maximal I_i^* .

Now, we examine contribution of each of the identified maximal subsets I_i^* to the LHS of eq. (9). It is evident that $|A_J| = 1$ for all $\emptyset \neq J \subseteq I_i^*$. Let $|I_i^*| = \delta$. This implies that the subset I_i^* induces the following contribution to the LHS of eq. (9): $\sum_{i=3}^{\delta} (-1)^{i+1} \binom{\delta}{i} = \binom{\delta}{0} - (\delta - 1) > 0$. Thus, the subset I_i^* of maximum cardinality contributes a net positive value to the LHS of eq. (9). Following this we can repeat this argument on the next maximal subset. Note that as the maximal subsets have an intersection of size at most one, each maximal subset contributes the LHS of eq. (9) via distinct terms. Finally, it can be observed that the overall contribution of the maximal subsets accounts for all terms in the LHS of eq. (9). We conclude that if there exist $|A_I| > 0$ for $\emptyset \neq I \subseteq [k]$, $|I| \geq 3$, we have $\sum_{\emptyset \neq I \subseteq [k], |I| \geq 3} (-1)^{|I|+1} |A_I| > 0$, which is a contradiction.

REFERENCES

- [1] R. Micheloni, A. Marelli, and K. Eshghi, *Inside Solid State Drives (SSDs)*. Springer, 2013.
- [2] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. on Info. Th.*, vol. 56, no. 9, pp. 4539–4551, 2010.
- [3] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. on Info. Th.*, vol. 58, no. 11, pp. 6925–6934, 2012.
- [4] D. S. Papailiopoulos and A. G. Dimakis, "Locally repairable codes," in *IEEE Intl. Symposium on Info. Th.*, 2012, pp. 2771–2775.
- [5] F. Oggier and A. Datta, "Self-repairing homomorphic codes for distributed storage systems," in *Proceedings IEEE INFOCOM*, 2011, pp. 1215–1223.
- [6] S. Jieka, A.-M. Kermarrec, N. L. Scouarnec, G. Straub, and A. V. Kempen, "Regenerating codes: A system perspective," *ACM SIGOPS Operating Systems Review*, vol. 47, no. 2, pp. 23–32, 2013.
- [7] S. E. Rouayheb and K. Ramchandran, "Fractional repetition codes for repair in distributed storage systems," in *48th Annual Allerton Conference on Communication, Control, and Computing*, 2010, pp. 1510–1517.
- [8] K. V. Rashmi, N. B. Shah, P. V. Kumar, and K. Ramchandran, "Explicit construction of optimal exact regenerating codes for distributed storage," in *47th Annual Allerton Conference on Communication, Control, and Computing*, 2009, pp. 1243–1249.
- [9] G. Kamath, N. Prakash, V. Lalitha, and P. Kumar, "Codes with local regeneration and erasure correction," *IEEE Trans. on Info. Th.*, vol. 60, no. 8, pp. 4637–4660, 2014.
- [10] C. J. Colbourn and J. H. Dinitz, *Handbook of combinatorial designs*. CRC press, 2010.
- [11] R. C. Bose, S. S. Shrikhande, and E. T. Parker, "Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture," *Canad. J. Math.*, vol. 12, pp. 189–203, 1960.
- [12] D. R. Stinson, *Combinatorial designs: construction and analysis*. Springer, 2004.
- [13] G. Exoo and R. Jajcay, "Dynamic cage survey," *The Electronic Journal of Combinatorics*, 2008.
- [14] O. Olmez and A. Ramamoorthy, "Repairable replication-based storage systems using resolvable designs," in *50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2012, pp. 1174–1181.
- [15] —, "Constructions of fractional repetition codes from combinatorial designs," in *47th Asilomar Conf. on Signals, Systems and Computers*, 2013, pp. 647–651.
- [16] —, "Replication based storage systems with local repair," in *International Symposium on Network Coding (NetCod)*, 2013, pp. 1–6.
- [17] K. V. Rashmi, N. B. Shah, and P. V. Kumar, "Optimal exact-regenerating codes for distributed storage at the MSR and MBR points via a product-matrix construction," *IEEE Trans. on Info. Th.*, vol. 57, no. 8, pp. 5227–5239, 2011.
- [18] C. Suh and K. Ramchandran, "Exact-Repair MDS Code Construction Using Interference Alignment," *IEEE Trans. on Info. Th.*, vol. 57, no. 3, pp. 1425–1442, 2011.
- [19] C. Tian, V. Aggarwal, and V. A. Vaishampayan, "Exact-repair regenerating codes via layered erasure correction and block designs," in *IEEE Intl. Symposium on Info. Th.*, 2013, pp. 1431–1435.
- [20] D. S. Papailiopoulos, J. Luo, A. G. Dimakis, C. Huang, and J. Li, "Simple regenerating codes: Network coding for cloud storage," in *Proceedings IEEE INFOCOM*, 2012, pp. 2801–2805.
- [21] N. B. Shah, K. V. Rashmi, P. V. Kumar, and K. Ramchandran, "Interference alignment in regenerating codes for distributed storage: necessity and code constructions," *IEEE Trans. on Info. Th.*, vol. 58, no. 4, pp. 2134–2158, 2012.
- [22] I. Tamo, Z. Wang, and J. Bruck, "MDS array codes with optimal rebuilding," in *IEEE Intl. Symposium on Info. Th.*, 2011, pp. 1240–1244.
- [23] D. S. Papailiopoulos, A. G. Dimakis, and V. R. Cadambe, "Repair optimal erasure codes through hadamard designs," in *49th Annual Allerton Conference on Communication, Control, and Computing*, 2011, pp. 1382–1389.
- [24] N. B. Shah, K. V. Rashmi, P. V. Kumar, and K. Ramchandran, "Distributed storage codes with repair-by-transfer and nonachievability of interior points on the storage-bandwidth tradeoff," *IEEE Trans. on Info. Th.*, vol. 58, no. 3, pp. 1837–1852, March 2012.
- [25] A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath, "Optimal locally repairable and secure codes for distributed storage systems," *IEEE Trans. on Info. Th.*, vol. 60, no. 1, pp. 212–236, 2014.
- [26] K. W. Shum and Y. Hu, "Functional-repair-by-transfer regenerating codes," in *IEEE Intl. Symposium on Info. Th.*, July 2012, pp. 1192–1196.
- [27] Y. Hu, P. P. C. Lee, and K. W. Shum, "Analysis and construction of functional regenerating codes with uncoded repair for distributed storage systems," in *Proceedings IEEE INFOCOM*, 2013, pp. 2355–2363.
- [28] J. C. Koo and J. T. Gill, "Scalable constructions of fractional repetition codes in distributed storage systems," in *49th Annual Allerton Conference on Communication, Control, and Computing*, 2011, pp. 1366–1373.
- [29] T. Ernvall, "The existence of fractional repetition codes," 2012, [Online] Available: <http://arxiv.org/abs/1201.3547>.
- [30] N. Silberstein and T. Etzion, "Optimal fractional repetition codes based on graphs and designs," *IEEE Trans. on Info. Th.*, vol. 61, no. 8, pp. 4164–4180, 2015.
- [31] G. M. Kamath, N. Silberstein, N. Prakash, A. S. Rawat, V. Lalitha, O. O. Koyluoglu, P. V. Kumar, and S. Vishwanath, "Explicit MBR all-symbol locality codes," in *IEEE Intl. Symposium on Info. Th.*, 2013, pp. 504–508.
- [32] G. Quattrocchi and H. Zeitler, "Hyperovals in steiner triple systems," *Journal of Geometry*, vol. 47, no. 1, pp. 125–130, 1993.
- [33] M. Greig and A. Rosa, "Maximal arcs in steiner systems $s(2, 4, v)$," *Discrete Mathematics*, vol. 267, no. 1, pp. 143–151, 2003.

- [34] E. F. Assmus, *Designs and their Codes*. Cambridge University Press, 1992.
- [35] T. Skolem, "Some Remarks on the Triple Systems of Steiner." *Mathematica Scandinavica*, vol. 6, pp. 273–280, 1958.
- [36] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.
- [37] F. Yates, "A new method of arranging variety trials involving a large number of varieties," *The Journal of Agricultural Science*, vol. 26, no. 03, pp. 424–455, 1936.
- [38] C. W. Lam, L. Thiel, and S. Swiercz, "The non-existence of finite projective planes of order 10," *Canad. J. Math*, vol. 41, no. 6, pp. 1117–1123, 1989.
- [39] H. Sachs, "Regular graphs with given girth and restricted circuits," *Journal of the London Mathematical Society*, vol. 1, no. 1, pp. 423–429, 1963.
- [40] F. Lazebnik and V. A. Ustimenko, "Explicit construction of graphs with an arbitrary large girth and of large size," *Discrete Applied Mathematics*, vol. 60, no. 1, pp. 275–284, 1995.
- [41] S. Jukna, *Extremal combinatorics*. Springer, 2001.
- [42] A. S. Tripathy and A. Ramamoorthy, "Capacity of sum-networks for different message alphabets," in *IEEE Intl. Symposium on Info. Th.*, 2015, pp. 606–610.

PLACE
PHOTO
HERE

Oktay Olmez received his Ph.D. in pure mathematics at the Iowa State University under the supervision of Dr. Sung Song in 2012. He also worked as a postdoctoral fellow in the Department of Mathematics and Department of Electrical and Computer Engineering at the Iowa State University between 2012 and 2013. He is currently an Associate Professor in the Department of Mathematics at Ankara University. His research interest include regenerating codes for distributed storage systems, highly regular graphs arising from finite geometries,

highly nonlinear boolean functions and construction of combinatorial block designs via difference sets.

PLACE
PHOTO
HERE

Aditya Ramamoorthy (M'05) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, in 1999, and the M.S. and Ph.D. degrees from the University of California, Los Angeles (UCLA), in 2002 and 2005, respectively. He was a systems engineer with Biomorph VLSI Inc. until 2001. From 2005 to 2006, he was with the Data Storage Signal Processing Group of Marvell Semiconductor Inc. Since fall 2006, he has been with the Electrical and Computer Engineering Department at Iowa State University,

Ames, IA 50011, USA. His research interests are in the areas of network information theory, channel coding and signal processing for bioinformatics and nanotechnology.

Dr. Ramamoorthy is the recipient of the 2012 Iowa State University's Early Career Engineering Faculty Research Award, the 2012 NSF CAREER award, and the Harpole-Pentair professorship in 2009 and 2010. He served as an associate editor for the IEEE Transactions on Communications from 2011 – 2014.